

Indo-European Language and Culture

An Introduction

Benjamin W. Fortson IV

© 2004 by Benjamin W. Fortson IV

BLACKWELL PUBLISHING

350 Main Street, Malden, MA 02148-5020, USA

9600 Garsington Road, Oxford OX4 2DQ, UK

550 Swanston Street, Carlton, Victoria 3053, Australia

The right of Benjamin W. Fortson IV to be identified as the Author of this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

First published 2004 by Blackwell Publishing Ltd

2 2005

Library of Congress Cataloging-in-Publication Data

Fortson, Benjamin W.

Indo-European language and culture : an introduction / Benjamin W. Fortson IV.
p. cm. — (Blackwell textbooks in linguistics ; 19)

Includes bibliographical references and index.

ISBN 1-4051-0315-9 (hardcover : alk. paper) — ISBN 1-4051-0316-7 (pbk. : alk. paper)

1. Indo-European languages. 2. Indo-Europeans. I. Title. II. Series.

P561.F67 2004

410—dc22

2004004901

ISBN-13: 978-1-4051-0315-2 (hardcover : alk. paper) — ISBN-13: 978-1-4051-0316-9 (pbk. : alk. paper)

A catalogue record for this title is available from the British Library.

Set in 10/12pt Sabon

by Graphicraft Limited, Hong Kong

Printed and bound in the United Kingdom

by TJ International Ltd, Padstow, Cornwall

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on

Blackwell Publishing, visit our website:

www.blackwellpublishing.com

Contents

List of Illustrations	xi
Preface	xii
Acknowledgments	xv
Guide to the Reader	xvii
1 Introduction: The Comparative Method and the Indo-European Family	1
The Study of Language Relationships and the Comparative Method (§§1.1–12)	1
Indo-European Historical Linguistics (§§1.13–18)	8
Conclusion (§§1.19–21)	12
For Further Reading	14
For Review	14
Exercises	15
2 Proto-Indo-European Culture and Archaeology	16
Introduction (§§2.1–2)	16
Society (§§2.3–16)	17
Religion, Ritual, and Myth (§§2.17–36)	22
Poetics (§§2.37–45)	29
Personal Names (§§2.46–49)	34
Archaeology and the PIE Homeland Question (§§2.50–73)	35
For Further Reading	44
For Review	45
Exercises	45
3 Proto-Indo-European Phonology	48
Introduction (§3.1)	48
Consonants (§§3.2–25)	48
Vowels (§§3.26–32)	60
Phonological Rules (§§3.33–44)	62
For Further Reading	66
For Review	66
Exercises	66

4	Proto-Indo-European Morphology: Introduction	69
	The Root and Indo-European Morphophonemics (§4.1)	69
	The Root (§§4.2–11)	70
	Ablaut (§§4.12–19)	73
	Morphological Categories of PIE (§§4.20–24)	76
	For Further Reading	78
	For Review	79
	Exercises	79
5	The Verb	81
	The Structure of the PIE Verb (§§5.1–10)	81
	Personal Endings (§§5.11–15)	83
	The Present Stem (§§5.16–40)	87
	The Aorist Stem (§§5.41–46)	92
	The Perfect Stem (§§5.47–49)	93
	Moods (§§5.50–53)	95
	Non-finite Verbal Formations and Other Topics (§§5.54–59)	96
	For Further Reading	99
	For Review	99
	Exercises	100
6	The Noun	102
	Introduction (§§6.1–3)	102
	Athematic Nouns (§§6.4–42)	103
	Thematic Nouns (§§6.43–67)	113
	The Collective and the Feminine (§§6.68–71)	118
	Adjectives (§§6.72–81)	120
	Nominal Composition and Other Topics (§§6.82–87)	122
	For Further Reading	123
	For Review	124
	Exercises	124
7	Pronouns and Other Parts of Speech	126
	Pronouns: Introduction (§7.1)	126
	Personal Pronouns (§§7.2–7)	126
	Other Pronouns and the Pronominal Declension (§§7.8–13)	129
	Numerals (§§7.14–21)	130
	Adverbs (§§7.22–24)	132
	Prepositions and Postpositions (§7.25)	133
	Conjunctions and Interjections (§§7.26–29)	134
	For Further Reading	135
	Exercises	136
8	Proto-Indo-European Syntax	137
	Introduction (§§8.1–5)	137
	Syntax of the Phrase (§§8.6–11)	139
	Syntax of the Clause (§§8.12–30)	141
	Phrase and Sentence Prosody and the Interaction of Syntax and Phonology (§§8.31–36)	149

For Further Reading	152
For Review	152
Exercises	152
9 Anatolian	154
Introduction (§§9.1–4)	154
From PIE to Common Anatolian (§§9.5–14)	156
Hittite (§§9.15–44)	158
Luvian (§§9.45–57)	167
Palaic (§§9.58–60)	171
Lycian (§§9.61–69)	172
Lydian (§§9.70–74)	175
Carian, Pisidian, and Sidetic (§9.75)	176
For Further Reading	177
For Review	177
Exercises	177
PIE Vocabulary I: Man, Woman, Kinship	178
10 Indo-Iranian I: Indic	180
Introduction to Indo-Iranian (§§10.1–4)	180
From PIE to Indo-Iranian (§§10.5–19)	181
Indic (Indo-Aryan) (§§10.20–22)	183
Sanskrit (§§10.23–51)	184
Middle Indic (§§10.52–57)	196
Modern (New) Indo-Aryan (§§10.58–63)	198
For Further Reading	199
For Review	200
Exercises	200
PIE Vocabulary II: Animals	202
11 Indo-Iranian II: Iranian	203
Introduction (§§11.1–8)	203
Avestan (§§11.9–25)	204
Old Persian (§§11.26–34)	210
Middle and Modern Iranian (§§11.35–51)	214
For Further Reading	218
For Review	219
Exercises	219
PIE Vocabulary III: Food and Agriculture	220
12 Greek	222
Introduction (§§12.1–8)	222
From PIE to Greek (§§12.9–53)	225
Greek after the Classical Period (§§12.54–57)	235
The Philology of Homer and Its Pitfalls (§§12.58–67)	237
For Further Reading	241
For Review	242
Exercises	242
PIE Vocabulary IV: The Body	244

13 Italic	245
Introduction (§§13.1–5)	245
From PIE to Italic (§§13.6–23)	248
Latino-Faliscan (§13.24)	252
Latin (§§13.25–53)	252
Faliscan (§§13.54–55)	262
Sabellic (Osco-Umbrian) (§§13.56–63)	263
Umbrian (§§13.64–71)	265
South Picene (§§13.72–73)	267
Oscan (§§13.74–77)	268
Other Sabellic Languages (§§13.78–79)	269
For Further Reading	269
For Review	270
Exercises	270
PIE Vocabulary V: Body Functions and States	273
14 Celtic	274
Introduction (§§14.1–3)	274
From PIE to Celtic (§§14.4–11)	275
Continental Celtic (§§14.12–19)	277
Insular Celtic (§§14.20–27)	280
Goidelic: Old Irish and Its Descendants (§§14.28–50)	282
Scottish Gaelic and Manx (§§14.51–52)	290
Brittonic (§§14.53–56)	291
Welsh (§§14.57–61)	292
Breton (§§14.62–68)	294
Cornish (§§14.69–72)	296
For Further Reading	297
For Review	298
Exercises	298
PIE Vocabulary VI: Natural Environment	299
15 Germanic	300
Introduction (§§15.1–4)	300
From PIE to Germanic (§§15.5–32)	301
Runic (§§15.33–36)	309
East Germanic (§15.37)	311
Gothic (§§15.38–45)	312
West Germanic (§§15.46–48)	315
Old English (§§15.49–61)	316
Middle and Modern English (§§15.62–66)	320
Old High German (§§15.67–76)	322
Old Saxon (§§15.77–80)	326
Dutch and Frisian (§§15.81–83)	327
North Germanic: Old Norse and Scandinavian (§§15.84–103)	328
For Further Reading	333
For Review	334
Exercises	334
PIE Vocabulary VII: Position and Motion	336

16 Armenian	337
Introduction (§§16.1–11)	337
From PIE to Classical Armenian (§§16.12–43)	340
For Further Reading	349
For Review	349
Exercises	349
PIE Vocabulary VIII: Material Culture and Technology	350
17 Tocharian	351
Introduction (§§17.1–6)	351
From PIE to Tocharian (§§17.7–33)	353
For Further Reading	362
For Review	362
Exercises	362
PIE Vocabulary IX: Form and Size	363
18 Balto-Slavic	364
Introduction (§18.1)	364
From PIE to Balto-Slavic (§§18.2–13)	365
Slavic (§§18.14–32)	368
Old Church Slavonic (§§18.33–35)	373
Modern Slavic Languages (§§18.36–48)	375
Baltic (§§18.49–60)	378
Lithuanian (§§18.61–65)	381
Latvian (§§18.66–67)	383
Old Prussian (§§18.68–70)	385
For Further Reading	387
For Review	387
Exercises	387
PIE Vocabulary X: Time	389
19 Albanian	390
Introduction (§§19.1–5)	390
From PIE to Albanian (§§19.6–26)	392
For Further Reading	398
Exercises	398
PIE Vocabulary XI: Utterance	399
20 Fragmentary Languages	400
Introduction (§§20.1–2)	400
Phrygian (§§20.3–9)	401
Thracian (§§20.10–11)	404
Macedonian (§20.12)	404
Illyrian (§§20.13–15)	405
Venetic (§§20.16–20)	406
Messapic (§§20.21–22)	407
Lusitanian (§20.23)	408
For Further Reading	409

--

Contents

Exercises	409
PIE Vocabulary XII: Basic Physical Acts	411
Glossary	412
Bibliography	416
Index	427

List of Illustrations

Figure 1.1	The Indo-European family tree	10
Map 1.1	Geographical distribution of the major Indo-European peoples around 500 BC	11
Map 2.1	Selected Late Neolithic and Chalcolithic (Bronze Age) cultures north of the Black and Caspian Seas	42
Map 9.1	Anatolia	155
Map 10.1	Modern Indo-Aryan languages	198
Map 11.1	The Persian Empire under Darius I	211
Map 11.2	Selected Modern Iranian languages	218
Map 12.1	Greek dialects	224
Map 13.1	Languages of ancient Italy	247
Map 14.1	The Celts	275
Map 15.1	The Germanic peoples around AD 500	311
Map 16.1	Greater Armenia during the early Christian era	339
Map 17.1	The Tarim Basin	352
Map 18.1	The Slavs and Balts around AD 1000	369
Map 19.1	Geographical distribution of Geg and Tosk	391
Map 20.1	The ancient Mediterranean	401

Preface

In spite of its venerable status as one of the oldest and most successful disciplines in linguistics, the comparative study of the Indo-European family has hitherto lacked an introductory textbook, or an introductory book of any kind appropriate for classroom use or for an intelligent layperson with linguistic interests but without specialized training. The few introductory books on the market do have virtues, but in all cases known to me these virtues are compromised, sometimes severely, as by idiosyncratic or minority views masquerading as *communis opinio*, by uneven coverage or omission of important topics, by out-of-date views or erroneous material, or by excessively technical information. None is a textbook.

Some of the features an Indo-European textbook should have are rather obvious. It should be up-to-date and, to the extent possible, present non-controversial views. It should not overwhelm the reader with detail, but also be comprehensive enough to satisfy the serious student at the beginning of his or her scholarly career. It should have copious exercises. Some other features an Indo-European textbook should have are less obvious (at least for some), but are in my view as essential as the preceding. It should not only cover phonology and morphology, but syntax as well, and incorporate relevant findings from generative linguistics where they are not limited to (or by) a particular (and quite likely evanescent) theoretical framework. Data should not be oversimplified or skewed by leaving out diacritics and other funny-looking symbols. There should be annotated text samples in all the ancient languages, and of a sufficient size to impart a real feel for the languages and to introduce readers to the practice and importance of philology. Basic information should be provided on the archaeological, cultural, and literary history of each branch. The modern languages should not be omitted from discussion. Finally, it should outline what we know about Proto-Indo-European culture and society.

These are the goals of the present work. If it falls short of any of them, hopefully it at least improves over other books currently available. It is designed for use with an instructor or for private consumption. It is assumed that the reader is interested in language and linguistic history, but no prior knowledge of linguistics or any older Indo-European language is necessary. Technical terminology is explained as needed, with a glossary appended for good measure.

Layout

Chapter 1 presents the tools of the trade and an overview of various basic issues that confront the researcher in the field. This is followed by a chapter on the

reconstructed culture of the Proto-Indo-Europeans, which attempts an organized and comprehensive introductory synthesis of a kind that I do not believe I have seen elsewhere. As interest in the matters discussed here usually runs high, this chapter has been put toward the beginning of the book rather than at the end.

Chapters 3–8 provide a reasonably complete introduction to reconstructed Proto-Indo-European that serves not only as background for the subsequent chapters, but also as a solid basic reference grammar in its own right. Not all the material included here will be needed in an introductory class; it is modular enough that individual instructors (and readers) can choose to omit whatever sections they please.

Each of the remaining twelve chapters treats a particular branch of the family, proceeding in chronological order of attestation from oldest to most recent and concluding with a chapter on the fragmentarily attested languages of uncertain filiation. Indo-Iranian is split over two chapters. All the branches and languages are treated as equally as possible, and in essentially the same format; but no attempt was made to make each chapter slavishly conform to identical specifications. In a book of this kind, no branch's history can or should be treated exhaustively. The developments discussed are selective; the chapters are tailored to what is interesting and important for each branch or language. Unlike the chapters on reconstructed Proto-Indo-European grammar, these are intended more to get one's feet wet than to give an overview that covers everything equally. The material can be supplemented at will by instructors, and again the modular structure of these chapters allows easy skipping of unneeded information.

Importantly, no language was deemed too trivial for coverage. The reader's interest should be engaged, and his or her curiosity piqued, with regard to *all* the Indo-European languages. The notion that certain languages are not particularly useful for Indo-European linguistics is both counterproductive and sadly perpetuated by books that only grant passing mention to those languages. Albanian may not, it is true, provide us with as much information about the proto-language as Sanskrit, but it does tell us some useful things, and in any event has a maddeningly fascinating (or fascinatingly maddening) history that merits careful attention. It bears repeating that Indo-European comparative linguistics is not just concerned with reconstructing Proto-Indo-European; it also must account for the histories of all the languages of the family.

In addition to the requisite basic sound laws, etc., short discussions of topics are included that are intended to appeal to readers with some experience in the languages, or to interested readers without any such background. The Greek chapter, for example, has a section on Homeric philology. A few topics that might seem a bit off the beaten path are scattered throughout. For instance, a student of first-year Russian, while in the throes of wrestling with the syntax of the cardinal numerals, may have become curious about how this unusual system came to be; a short discussion of this in the section on Slavic is therefore included.

The Sample Texts and Other Features

The text samples are meant to give an idea of what the languages look like, and to impress the student with the importance of philology, whose methods and purpose are outlined in chapter 1. The texts are about a paragraph in length on average. It is to be hoped that exposure to good-sized text samples with philological commentary

will inspire an interest not only in the languages themselves, but also in close textual and etymological analysis. Where possible I have chosen texts having literary and cultural interest; many are connected to discussions in chapter 2. The commentary is geared especially to the reader interested in learning a bit more: it provides word histories and comparanda, points out examples of developments covered in the chapter, and also adds notes on developments not met with in the chapter. Unless indicated otherwise, all translations are my own.

I decided to eschew the usual interlinear word-by-word analytical glosses that are otherwise standard in the linguistic literature; the commentary stands partly in lieu of them. The reason for doing this is perhaps unorthodox, but I believe sound: interlinear glosses distort one's perception of the aesthetics of a written language, which is not a trivial concern for a book such as this. The look of a language should never be underestimated as a tool for engaging a student's interest. The commentary at any rate usually makes it possible for students who seriously want to match each word of the original up with its translation to do so. (For some longer passages, however, and in some longer chapters, commentary could only be given selectively, due to limitations of space.)

This leads to one further point. Some who have written introductory materials on Indo-European have left out phonetic details, such as long marks and accents. Such simplification of the data renders all students a great disservice by selling the languages short. Someone likely to pick up a book such as this or to take an introductory Indo-European course will not be put off by unusual marks and symbols. Quite the opposite – they are likely to be *intrigued* by the peculiar look of the strange forms, and will discover in them an inviting mystery and beauty. The decision to ignore “details” such as accents should at any rate be left to the reader's discretion.

Each chapter closes with several additional sections. The “For Further Reading” sections provide brief commentary on the most important or prominent secondary literature; the full references (together with a few extra that are not discussed) are listed in the Bibliography preceding the Index. Devoting space to such commentary rather than to long bibliographical lists is more useful for this kind of book. The “For Review” sections list the main terms and concepts, and the “Exercises” that follow are designed both for review and for going beyond the material discussed in each chapter. Finally, starting in chapter 9, a short list of reconstructed roots or words in Proto-Indo-European is given, arranged by semantic category. How instructors choose to integrate these in with their course is left up to them.

Since few of the localities mentioned in the text will be familiar to most readers, illustrative maps are provided in each chapter on the branches as well as chapters 1 and 2. These were deemed to be the most important visual component of the book. Chapter 1 also contains a diagram of the Indo-European family tree. Regrettably, cost and space limitations prohibited the inclusion of other planned illustrations, as of artifacts and ancient scripts; it is hoped that these can appear in a future edition.

Acknowledgments

I have received the invaluable assistance of many colleagues, students, and friends since I conceived the idea of writing this textbook half a decade ago. Jay Jasanoff and Brent Vine used earlier drafts in their introductory Indo-European courses at Harvard and UCLA and provided extraordinarily helpful and detailed comments, corrections, and suggestions on the entire manuscript. I deeply appreciate the time and effort they both gave so generously to making this a better book.

I have been extremely fortunate to benefit from many others' interest and input as well. Raimo Anttila, Lisi Oliver, and Timothy Pulju also used earlier drafts in their courses at UCLA, LSU, and Dartmouth, respectively. Tim read through the whole manuscript and saved me from quite a few mistakes. Lisi engaged her students in writing reviews of the book, which generated much useful input at an early stage and guided me to a better presentation of the material. My own students over several years of teaching, especially in my Indo-European course at the Harvard Extension School, caught errors, made important suggestions, and – perhaps most importantly – convinced me that the project was both workable and worthwhile, and that my organizational and presentational decisions were on the whole satisfactory. I would like to single out in particular Valerie Goodspeed, Matthew Harrington, William Heess, Ilya Lapshin, Robert Maher, Keith Plaster, and William Waters for their helpful comments and criticisms.

Many other colleagues and friends improved the book by their careful reading and valuable input, and provided very helpful critiques of individual chapters or larger portions of the work: Kelly Askew, Don Cameron, Steve Dworkin, Simon Eckley, Joseph Eska, Michael Flier, Mark Hale, Joseph Harris, Richard Janko, Robert Kyes, J. P. Mallory, H. Craig Melchert, Alan Nussbaum, Hugh Olmsted, Marc Pierce, P. Oktor Skjærvø, Bert Vaux, Aurelijus Vijūnas, Rex Wallace, and Calvert Watkins. To Cal I also owe some innovative organizational suggestions that I believe have enhanced the book's readability and interest. All of the people above not only provided material contributions in the form of corrections and so forth, but were unflagging in their encouragement and enthusiasm for this project. Writing this textbook has mostly been its own reward, but it has meant a great deal that so many people that I look up to have been so supportive and positive in their assessment. At the University of Michigan I was lucky enough to have a recently graduated Senior, Dan Brooks, volunteer to read the whole manuscript, work through all the exercises, and assist me in proofreading and fact-checking; his keen attention caught all manner of infelicities that had escaped others. I was able to pay Dan for

his services with a generous subvention from the Block Funding Initiative of the College of Literature, Science, and the Arts at the University of Michigan.

Had I not been initially contacted by Tami Kaplan (then at Blackwell) and encouraged to submit a very rough-and-ready chapter of the manuscript, I am sure the book would still lie unfinished and unpublished to this day. I have Tami to thank for seeing it through the approval process, and Sarah Coleman, my editor at Blackwell, receives my gratitude for so ably overseeing the book's progression from submitted manuscript to published volume. In a day when computerized typesetting has too often fallen short of the exacting aesthetic requirements of good Indo-European publications, I wish to shine a most grateful spotlight on the superb typographical skills of Graphicraft Limited, whose eye-winning work graces every page. To Chartwell Illustrators and their careful attention to cartographic detail I owe the conversion of even my most involved instructions into all the highly readable maps in the book. And without the careful copy-editing by Margaret Aherne, numerous errors and inconsistencies would still remain in every chapter; to her I am deeply indebted for her diligence, eagle eyes, and patience.

My family has given me great encouragement throughout this long process. I thank my mother for her unflagging support and positive remarks on the manuscript, and my father for his enthusiasm and infectious eagerness to see the finished product. For all the happiness and support that Kelly and Christopher have brought me at every stage of this work, and for so much else besides, I am more grateful than words can express.

As often as I could, I took the advice of all these good people I have mentioned, and am alone responsible for the glee or frustration that students will have at finding whatever errors and infelicities remain.

Guide to the Reader

A. Abbreviations

abl.	ablative	Luv.	Luvian
accus., acc.	accusative	masc.	masculine
act.	active	ME	Middle English
adj.	adjective	Mod.	Modern
Alb.	Albanian	neut.	neuter
Anat.	Anatolian	nomin., nom.	nominative
aor.	aorist	OAv.	Old Avestan
Arm.	Armenian	OCS	Old Church Slavonic
Av.	Avestan	OE	Old English
B.-Sl.	Balto-Slavic	OHG	Old High German
Celt.	Celtic	OHitt.	Old Hittite
Class.	Classical	Oldr.	Old Irish
Cz.	Czech	OLith.	Old Lithuanian
dar.	dative	ON	Old Norse
du.	dual	OPers.	Old Persian
Du.	Dutch	OPruss.	Old Prussian
Eng.	English	OS	Old Saxon
fem.	feminine	Osc.	Oscan
fut.	future	pass.	passive
Gaul.	Gaulish	PIE	Proto-Indo-European
genit., gen.	genitive	pl.	plural
Gk.	Greek	Pol.	Polish
Gmc.	Germanic	pres.	present
Goth.	Gothic	Russ.	Russian
Hitt.	Hittite	sing., sg.	singular
Icel.	Icelandic	Skt.	Sanskrit
IE	Indo-European	Slav.	Slavic
imperf.	imperfect	Toch.	Tocharian
Indo-Ir.	Indo-Iranian	Umbr.	Umbrian
instr.	instrumental	Ved.	Vedic
Lar.	Latin	voc.	vocative
Latv.	Latvian	W.	Welsh
Lith.	Lithuanian	YAv.	Younger Avestan
loc.	locative		

B. Symbols

- * denotes a reconstructed form, not preserved in any written documents
- < “comes from” or “is derived from”
- > “turns into” or “becomes”
- indicates morpheme boundary, or separates off that part of a word that the reader should focus on
- () encloses part of a word that is not relevant to the discussion, or that is an optional part
- ~ separates pairs of examples or forms

C. Spelling Conventions

All linguistic forms are written in *italics*. The only exceptions are inscriptional forms in Italic dialects (such as Oscan and Umbrian) that are not written in the Latin alphabet; these, following standard convention, are given in **boldface**. See chapter 13. For Latin, *i* and *u* are used for both the vowels and the glides (instead of *j* and *v*).

D. International Phonetic Alphabet (IPA)

Phonetic transcriptions using the IPA are enclosed in square brackets. The symbols used for American English sounds are:

Consonants

b	bell	p	pat
d	dim	r	roof
ð	this	s	silver
f	fail	ʃ	shelf
g	go	t	tin
h	heal	θ	thin
j	yarn	v	vat
k	coal	w	well
l	light	z	zero
m	magic	ʒ	measure
n	near		
ŋ	sing		

Vowels

a	father
æ	hat
e	care
ɛ	pet
ə	about
i	beat
ɪ	bit
o	bore
ɔ	bought
u	boot
ʊ	book
ʌ	but

1 Introduction: The Comparative Method and the Indo-European Family

The Study of Language Relationships and the Comparative Method

1.1. All languages are similar in certain ways, but some similarities are more striking and interesting than others. Consonants, vowels, words, phrases, sentences, and their ilk are fundamental structural units common to all forms of human speech; by contrast, identical or near-identical words for the same concept are not, and when two or more languages share such words, it attracts notice. This kind of resemblance can have several sources, which must be clearly distinguished from one another in order to investigate similarities between languages scientifically.

The first source for such resemblance is **chance**. There are only so many sounds that the human vocal tract can produce, and their possible combinations are also limited. These facts conspire to create a certain number of words that coincidentally resemble one another in any two languages picked at random. The Greek and Latin words for 'god', *theós* and *deus*, are of this kind; they have no historical relationship with one another.

A second source of such similarity is **borrowing**. People speaking different languages are often in contact with one another, and this contact typically leads to mutual borrowing (adoption) of both cultural and linguistic material. English, for example, has borrowed the Inuit (Eskimo) word *iglu* 'house' for a type of shelter (*igloo*).

A third source of similarity is a **sundry collection of language universals**; these are basic characteristics of human linguistic creativity that are found the world over. Two common examples are onomatopoeia or sound-symbolism (whereby words sound like what they mean, such as English *cuckoo* and Germanic *Kuckuck*, names based on imitation of the bird's cry), and nursery or baby-talk words for kinship terms, which typically contain syllables like *ma*, *ba*, *da*, and *ta* (compare English *Ma* with Mandarin Chinese *mā* 'mother').

1.2. Sometimes, however, languages present similarities in their vocabulary that cannot be attributed to any of these sources. To take a concrete example, consider the words for the numerals 1–10 in Spanish, Italian, French, and Portuguese:

	Spanish	Italian	French	Portuguese
1	uno	uno	un	um
2	dos	due	deux	dois
3	tres	tre	trois	três
4	cuatro	quattro	quatre	quatro
5	cinco	cinque	cinq	cinco
6	seis	sei	six	seis
7	siete	sette	sept	sete
8	ocho	otto	huit	oito
9	nueve	nove	neuf	nove
10	diez	dieci	dix	dez

The striking similarities in each row attract immediate attention and demand an explanation. Chance seems well-nigh impossible. There is also no connection between the sounds of these words and their meanings, which rules out onomatopoeia; nor are other linguistic universals such as baby-talk relevant. A third possibility is that one or more of the languages borrowed its numerals from one of the other languages, or that they all borrowed them from some outside source. It is true that there are languages that have borrowed the names of some or all numbers from other languages, as Japanese did from Chinese. But if we look a bit further afield, we notice that the numerals are not the only words evincing such strong mutual resemblance:

	Spanish	Italian	French	Portuguese
(1) 'two'	dos	due	deux	dois
'ten'	diez	dieci	dix	dez
'tooth'	diente	dente	dent	dente
'of'	de	di	de	de
'they sleep'	duermen	dormono	dorment	dormem
(2) 'am'	soy	sono	suis	sou
'you (sing.) are'	eres	sei	es	és
'is'	es	è	est	é
'we are'	somos	siamo	sommes	somos
'you (pl.) are'	sois	siete	êtes	sois
'they are'	son	sono	son	são

We see that not just the words for 'two' and 'ten', but all the other words in group (1) above agree in beginning with *d-* in each language. It is rather uncommon for basic terms like 'tooth', 'of', and 'sleep' to be borrowed, and even more uncommon for three languages to have borrowed them from a fourth, or for all four of them to have borrowed these words from a fifth language. The forms in (2) show that the whole present tense of the verb 'to be' is similar across all four languages, and in very specific ways. It is extremely unlikely that a language (to say nothing of four languages) would borrow wholesale a complete verbal paradigm from another language, especially one as basic as this one – and one that, as it happens, is highly irregular in nearly the same way in each language.

1.3. If two or more languages share similarities that are so numerous and systematic that they cannot be ascribed to chance, borrowing, or linguistic universals, then the only hypothesis that provides a satisfactory explanation for those similarities is that they are descended from the same parent language. This is the essential statement of what is known as the **comparative method**. And in the case of Spanish, Italian, French, and Portuguese this hypothesis would be right: we know from other evidence that these languages are all descended from a variety of Latin.

Languages like these that are descended from a common ancestor are said to be **genetically related**. This technical term has nothing to do with biology; it makes no claims about the race or ancestry of the *speakers* of the languages in question, who may belong to many ethnicities. (Just think of all the different ethnic backgrounds of people speaking English as their native language within any large English-speaking city.)

Comparative reconstruction

1.4. So far we have shown how genetic relationship can be demonstrated, and in one sense our task – of explaining the similarities between Spanish, Italian, French and Portuguese – is done. But historical and comparative linguists typically do not stop there; they are also interested in figuring out what a putative ancestral language was like – in other words, to reconstruct it. Reconstruction is accomplished through systematic comparison of the forms in the descendant languages. Here we must content ourselves with one brief illustration. Let us take another look at the words for ‘tooth’ in our four languages above:

Spanish	Italian	French	Portuguese
diente	dente	dent	dente

All these words begin with *d-*, meaning their ancestor surely began with *d-* as well. The four words also agree in having the consonant cluster *nt* (ignore for the moment the fact that the *nt* in the French form is not pronounced as *nt*; in older French it was pronounced as written). The ancestral word thus probably had the “skeleton” *d . . . nt . . .* Italian, French, and Portuguese agree in having *e* before the consonant cluster, but Spanish has a diphthong *ie*. Since only Spanish is deviant here, the simplest thing (barring evidence to the contrary) is to suppose that it changed an earlier *e* to *ie*, rather than that the other three each changed an earlier *ie* to *e*. In fact, if we were to look at other examples, we would find that this was a regular sound change in the history of Spanish. Finally, all the languages except French agree in having the word end with the vowel *e*; since in the general course of language change sounds are not added to words willy-nilly, we may suppose that French lost an original final vowel here that is still preserved in the other three languages. Final-vowel loss of this kind is in fact extremely common cross-linguistically.

Putting all this information together, we may surmise that the ancestral word for ‘tooth’ had the shape *dente*. As a final but crucial touch we must add an asterisk before this reconstruction (**dente*), which is the conventional marker in historical linguistics for a hypothetical form – one that is not actually attested (preserved in documents) but is thought to have once existed. (As some readers may know, asterisks have other uses in other branches of linguistics, such as to denote ungrammatical sentences; the historical linguistic usage should not be confused with those.)

1.5. Each of the groups of words that we have been comparing with one another are called **correspondence sets**, and the words in each correspondence set are termed **cognates**. Thus the Spanish cognate of French *dent* is *diente*, the Italian cognate of French *dix* is *dieci*, etc. While we based an example of reconstruction on a sole correspondence set, in actual practice many correspondence sets must be examined for reconstructions to have much weight. Sound correspondences across one set of cognates must recur in other sets to be of any scientific worth. This principle is known as the **regularity of sound correspondences**; without it, comparative linguistics would be impossible.

1.6. An ancestral language is called a **proto-language**, and its descendants are termed its **daughter languages**. In this book, the phrase “the proto-language” will refer to Proto-Indo-European (see below). The prefix *proto-* is attached to the name of a language or group of languages to designate the immediate ancestor of that language or group; one therefore speaks of *Proto-Greek*, *Proto-Germanic*, *Proto-Chinese*, etc. The word for ‘tooth’ that we reconstructed above is in a language we can call Proto-Romance, the common ancestor of the Romance languages. (Note in passing that Proto-Romance was not the same as standard Classical Latin, where the word for ‘tooth’ was *dēns* [stem *dent-*].)

Importantly, observe that terms like *proto-language* and *Proto-Romance* do not designate a “prototype” language that still needed some time in the shop before it could be billed a “real” language. A central finding of linguistics has been that all languages, both ancient and modern, spoken by both “primitive” and “advanced” societies, are equally complex in their structure. We have no reason to believe that reconstructed, unattested languages were qualitatively any different from attested ones: the ability to speak complex language is common to all members of the species *Homo sapiens*, and that species has changed little if at all over the past 100,000 years.

Language change

1.7. The fact that a single language (such as Latin) can develop into two or more different languages (such as Spanish or French) is due to language change. We cannot here present a detailed discussion of the causes of language change, but will outline just a few basic points, a bit simplified for brevity. Linguists view language as a cognitive faculty whose core structures mostly develop during the first few years of childhood. A child’s native language is acquired from scratch. Contrary to popular wisdom, no one teaches children their native language; they must analyze the speech of people around them and construct their own individual **grammar** of the language. A grammar, in linguistic parlance, is a body of knowledge consisting of unconscious rules and principles; it may be conceived as the invisible underlying machinery used to produce and comprehend linguistic utterances in a particular language. No one has direct access to anyone else’s grammar, only to speech – the output of a grammar; because of this, the new grammar that the child constructs may well turn out to be subtly different from the grammars of his or her parents and other people in the child’s environment. Any of these differences can be called **linguistic changes**, and can in principle be picked up by other speakers and spread through part or all of a speech community.

Over successive generations, these differences multiply; as some of them diffuse throughout a community, the speech of the community evolves differently from that of other communities. We may then talk of communities that have developed their

own *dialects*; and given enough time these dialects can develop into what eventually may be labelled different *languages*. (These terms are not scientific, but useful for general descriptive purposes.)

1.8. Language change, whatever its precise mechanism, is an entirely natural phenomenon, part and parcel of every living language. A number of popularly held misconceptions may cloud appreciation of this point. It is often claimed that languages change to become easier or simpler. But as was said above, all human languages, past and present, exhibit the same level of complexity; they are also learned by children at the same rate, with equal ease, and in the same well-defined developmental stages. Linguistic difficulty is a purely subjective valuation, and has no basis in scientific fact. Another common belief is that languages change through laziness, ignorance, stupidity, or some benighted combination thereof. This view is typically taken by those for whom a particular linguistic stage or style is perfect and sublime, and all subsequent deviations from it are a product of decay. But this "sublime" form of the language is itself always a "decayed" development of an earlier stage. The changes happening today that are so often decried are in fact no different from the changes that languages have always undergone.

1.9. When the sounds of a language change, one speaks of *sound change*. Sound changes, importantly, are *regular* and *exceptionless* – that is, they affect all the relevant examples of the particular sound(s) in the language. This claim about sound change is called the *Neogrammarian hypothesis*, named after the Neogrammarians, an influential group of nineteenth-century linguists. A sound change in a language that turns a *p* between vowels into *b*, say, will change every intervocalic *p* in the language to *b*. The regularity of sound change accounts for the regularity of sound correspondences between related languages that we discussed above. Tracing sound change is often easier than tracing other kinds of change, and for this reason a listing of sound changes makes up a large and detailed portion of the historical sketches in this book. A fairly extensive terminology has been developed to label the various results of sound change; these and other technical terms will be defined as they arise and are also listed in the Glossary.

All other components of language change as well. The smallest linguistic units that have meaning are *morphemes* (whole words such as *foot*, *oyster*, *devil*, or prefixes and suffixes like *un-* or *-ing*); the rules for using and combining morphemes constitute the *morphology* of a language, and any change to these rules is called *morphological change*. A common type of this is change in the *productivity* of a morpheme – that is, in how freely it can be used to form new words or grammatical forms. The *-th* in words like *sloth*, *breadth*, and *filth* was once a productive suffix for forming nouns from adjectives (*slow*, *broad*, and *foul*, respectively), but is no longer (its function has been taken over by *-ness* and other suffixes); a reverse development is illustrated by the English plural suffix *-s*, now limitlessly productive but once used only with certain classes of nouns. Words and morphemes are stored in the *lexicon*, one's mental dictionary; changes to individual words (rather than to whole classes of words at once, as when a morphological rule has changed) constitute *lexical change*. The replacement of the old plural *kine* by *cows*, and of the old past tense *holp* by *helped*, are examples of a lexical change; in cases such as these, an old irregular form (containing morphemes or morphological processes that are no longer productive) is replaced by a regular form, by a process called *analogy*. The words stored in the lexicon are combined into larger units (phrases, clauses, sentences) by rules encoded in a language's *syntax*. Change to these rules is *syntactic change*, as

when a language that used to put verbs at the ends of clauses now puts them at the beginning. Finally, the meanings of words, also stored in the lexicon, constitute the words' semantics; changes to word meanings constitute *semantic change*. This is really a subtype of lexical change, since only individual words are affected; but sometimes there are far-reaching ramifications of semantic change, as when a noun like French *pas* 'step' gets specialized as a grammatical marker (in this case the negator, 'not'), by a process called *grammaticalization*.

Determining the pronunciation of dead languages

1.10. The ability to compare ancient languages rests upon knowledge of the sounds and grammatical structures of those languages. How does one figure out such facts about languages that are no longer spoken? The basic answer is that one uses everything that is at one's disposal: contemporary descriptions, the testimony of descendant languages, orthographic practice (including spelling errors), the rendering of loanwords from known source languages, and metrical evidence from poetry.

Sometimes, as in the case of Sanskrit, we are fortunate in having detailed descriptions of the language's pronunciation and structure by ancient grammarians. If a language has living descendants, as in the case of Latin, we can apply the comparative method to the descendants to establish the pronunciation of their ancestor. When we lack thorough contemporary descriptions or the testimony of living descendants, we must look to texts as they were written by speakers of the language.

Misspellings are valuable for revealing the effects of changes in pronunciation. We know, for example, that in the late pre-Christian era, the Latin diphthong spelled *ae* came to be pronounced as a monophthong (single vowel) *e*. A Roman in the first century AD who spelled the word *aetate* 'age' improperly as *etate* shows the effects of this change. Standard spelling conventions, not just misspellings, can also elucidate facts about pronunciation. In Hittite, like English, words were written with spaces between them; but various short function words like *ma* 'but' were joined to a preceding word without a space, indicating that they and the preceding word were pronounced together as a unit. (Compare the *-n't* of English *didn't* for a similar situation.) Such spelling conventions will be discussed further in §8.33.

We can glean further information from the way words are spelled when borrowed from another language. Educated Romans of the first century BC regularly rendered the Greek letter phi (Φ) as *ph* in words that they borrowed from that language (as in *philosophia* 'philosophy'). The sequence *ph* was not used in writing native Latin words, so the Romans must have been trying to represent a sound that they did not have in their own language – an aspirated stop consonant, as it happens, different from the native Latin unaspirated *p* and also from the native Latin fricative *f*. A few centuries later, however, we find the Gothic bishop Ulfilas using *f* (as in the name *Flippus*), showing that the pronunciation of the letter in Greek had changed.

Poetry is often a useful source of information on how a language is pronounced. Much ancient poetry is structured according to particular sequences of heavy and light syllables. In the Archaic Latin poetry of Plautus, a word like *patre* 'by the father' scans as two light syllables. Light syllables end by definition in a short vowel, so the word was pronounced *pa-tre* with the consonant cluster *tr* not split between the two syllables. Metrical practice has thus revealed a fact about Latin syllabification.

1.11. With regards to larger levels of linguistic structure such as morphology and syntax, and with regards to word-usage and meaning, we must always look to the original texts. Linguistic analysis of texts is the domain of **philology** – an enterprise that has been neatly summarized as “the art of reading slowly.” It is a multifaceted discipline with many uses, including the determination of the original wording of texts by comparing and dating extant manuscripts, as well as the determination of grammatical facts by examining how forms are used in context and paying attention to relative ages of particular usages. (“Philology” is also a somewhat old-fashioned term for what we now call comparative historical linguistics, hence the phrase “comparative philology”).

Both these and other aspects of philology are crucial for writing an accurate history of a language’s development. When texts are copied and recopied by hand and original source manuscripts (called *archetypes*) are lost, errors and modernizations inevitably creep in. If the language of the text is particularly archaic compared with the language of the copyist, errors can become legion. Careful comparison of different versions of the same text usually allows one to reconstruct chains of transmission and to determine the relative ages of manuscripts, which in turn helps us determine which linguistic forms are older and which more recent.

1.12. Using philology to figure out how words and grammatical forms are used has allowed countless puzzling details about ancient languages to be clarified. Dictionaries, being essentially lists of words, do not give a whole picture of a language, much less its history, and the forms they contain cannot always be taken at face value. Two examples from ancient Greek and Sanskrit will make this point clear.

The first example concerns the Greek word *prophētēs* ‘prophet’. This word etymologically consists of the combination form *-phētēs* ‘sayer’, derived from the verb *phēmi* ‘I say’, preceded by the prefix *pro-*, which can mean either ‘before’ or ‘forth’. Does the word mean ‘one who says beforehand, one who foretells’ or ‘one who speaks forth, one who announces’? There is a verb *prōphēmi* ‘I speak before’, and so we might jump to the conclusion that *prophētēs* comes from this verb and means ‘one who speaks beforehand’. But a look at the actual textual occurrences of these words shows that *prophētēs* first appears a good 700 years *before* the earliest attestations of *prōphēmi*. Clearly the noun cannot be derived from a verb that did not yet exist. Additionally, an investigation of the other compounds beginning with *pro-* reveals that the meaning ‘before’ is not the prefix’s oldest meaning, and does not appear until later than the first occurrences of *prophētēs*. We conclude that a ‘prophet’ was originally one who ‘spoke forth’ or ‘announced’ the will of the gods rather than one who foretold the future.

The second example concerns the Sanskrit word *nāveda-* ‘knowledgeable’. This word is clearly related to the word *vēda-* ‘knowledge’, but the first element, *na-*, is not a known prefix elsewhere in the language, and it seems to have little effect on the overall sense. An examination of the earliest passages in which the word occurs, in the Rig Veda (the oldest Sanskrit text), suggests an explanation for this form. Frequently, *nāveda-* is part of a phrase meaning ‘be knowledgeable (of)’. One version of this phrase is *bhūta nāvedāḥ* ‘be ye knowledgeable’, where the word *bhūta* is the plural imperative of the second person of the verb ‘to be’. A glance at a Vedic grammar-book will show that there is a common alternate version of this form, namely *bhūtana*, with an element *-na* that can optionally be added to the second plural. We can thus reasonably surmise that someone, in the course of the

transmission of the text, misanalyzed this particle as a prefix on the following word, thereby creating the form *náveda-* which then spread to other versions of this phrase. (A similar sort of false division has happened many times in English, as when the older phrase *an ekename* 'a supplementary name' was reanalyzed as a *nekename*, the source of our word *nickname*.)

Indo-European Historical Linguistics

1.13. Already in classical antiquity, it was noticed that Greek and Latin bore some striking similarities to one another like those that we saw among the Romance languages. Ancient writers pointed out, for example, that Greek *héks* 'six' and *heptá* 'seven' bore a similarity to Latin *sex* and *septem*, even pointing out the regular correspondence of initial *h-* in Greek to initial *s-* in Latin. The ancients explained such facts by viewing Latin as a descendant of Greek. During and after the Renaissance, as the vernacular languages of Europe came to be known to scholars, it slowly became understood that certain groups of languages were related, such as Icelandic and English, and that the Romance languages were derived from Latin. But no consistent scientific approach to language relationships had been developed.

1.14. Following the British colonial expansion into India, a language came to the attention of Western scholars knowledgeable in Greek and Latin that ushered in a new way of thinking about such matters. An orientalist and jurist named Sir William Jones was the first to state this way of thinking, in a lecture to the Asiatick Society on February 2, 1786, and published two years later in *Asiatick Researches* 1:

The Sanscrit language, whatever be its antiquity, is of a wonderful structure; more perfect than the Greek, more copious than the Latin, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of verbs and in the forms of grammar, than could possibly have been produced by accident; so strong indeed, that no philologist could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists: there is a similar reason, though not quite so forcible, for supposing that both the Gothic and the Celtic, though blended with a very different idiom, had the same origin with the Sanscrit; and the old Persian might be added to the same family, if this were the place for discussing any question concerning the antiquities of Persia.

This was a turning point in the history of science. For the first time the idea was put forth that Latin was not derived from Greek, but that they were both "sisters" (as we would now call them) of each other, derived from a common ancestor no longer spoken. The idea was inspired by the critical discovery of the third member of the comparison (the *tertium comparationis* in technical jargon), namely Sanskrit – a language geographically far removed from the other two. Also, this passage contains the first clear formulation of the central principle of the comparative method.

The IE family: branches, subgrouping, models

1.15. Jones's insight marks the beginning of the scientific study of the language family now called Indo-European, or IE for short. The field is variously known as Indo-European historical linguistics, Indo-European comparative linguistics, or Indo-European comparative philology. Jones's brief statement already enumerates fully

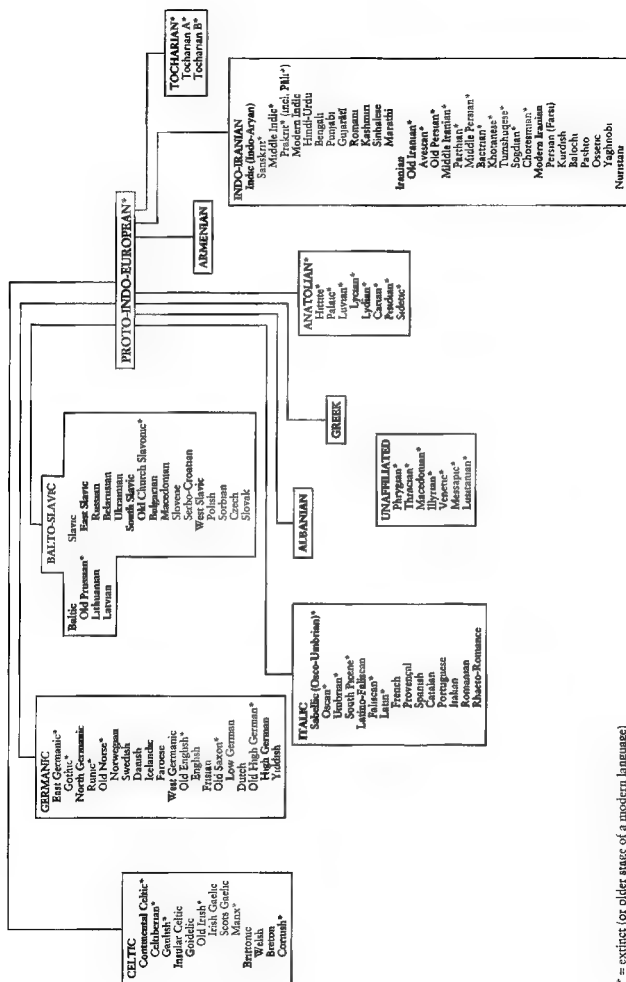
half the branches now recognized for the family: **Indo-Iranian** (containing Sanskrit and the Iranian languages), **Greek, Italic** (containing Latin and related languages of Italy), **Celtic**, and **Germanic** (containing Jones's "Gothick" and its relatives of northern Europe, including English).

After Jones's pronouncement, nearly three decades passed before any activity that could be called Indo-European linguistics arrived on the scene. Once it did arise it came rather fast and furious, with three seminal works published within a half-dozen years in Denmark and Germany, including one by Jacob Grimm, one of the two Grimm brothers of fairy-tale-collecting fame. These pioneers and the scholars who followed them realized that not only the languages in Sir William's account were "sprung from some common source," but also the Baltic and Slavic languages (now grouped together as Balto-Slavic), **Armenian**, and **Albanian**. In the twentieth century two more branches, **Anatolian** and **Tocharian**, were added to the family, containing extinct languages only discovered in the early 1900s. Anatolian, it turned out, was the most ancient of them all, with texts in Hittite dating to the early or mid-second millennium B.C. A few other languages with only meager remains are also clearly Indo-European, such as Phrygian, Thracian, and Messapic; whether they belong to any of the ten recognized branches, or constitute separate branches of their own, is not clear.

The nineteenth century also saw the creation of a name for the family, Indo-European or, in German-speaking lands, usually *Indogermanisch* 'Indo-Germanic'. The ancestor of all the IE languages is called *Proto-Indo-European*, or PIE for short. During the course of the nineteenth century, the methods of comparative linguistics and linguistic reconstruction were developed; a prodigious number of important discoveries in the watershed decade of the 1870s was responsible for significant refinement of the method and for firmly establishing historical linguistics as a science unto its own. By the dawn of the twentieth century, a picture of reconstructed PIE had emerged that was quite similar to the one that is presented in this textbook.

All standard diagrams of the IE family tree, such as the one in figure 1.1, represent it as a starburst: PIE is placed at the top or center, with the ten branches radiating out therefrom. Implicit in the traditional diagrams is the notion of a more or less simultaneous "breakup" of the proto-language into ten or more dialect areas (the future branches). This notion has often come under attack. For one thing, the trees of most other language families tend to have a binary branching structure. In addition, it has been speculated for a long time that certain branches of IE are to be grouped together into what are known as *subgroups*. For example, it has been proposed that Indo-Iranian, Greek, and Armenian all descend from one common dialect area of late PIE; and the same has been forwarded for Italic and Celtic, for which an "Italo-Celtic" subgroup has been suggested.

1.16. The discovery of Anatolian and Tocharian in the twentieth century has further fueled the debate over the internal structure of the PIE family tree. Anatolian in particular is significantly different from the picture of PIE that scholars had developed by the close of the nineteenth century on the basis of the other branches, even though Anatolian is the oldest attested branch. In the 1930s, the American linguist Edgar Sturtevant proposed that PIE was not the ancestor of Anatolian, but a sister of it, and that both PIE and Anatolian were descended from a language he called Indo-Hittite. Sturtevant's theory was not widely followed, and the term "Indo-Hittite" has largely been abandoned. But the hypothesis that Anatolian, and



* = extinct (or older stage of a modern language)

Figure 1.1 The Indo-European family tree, showing the approximate geographical distribution of the branches and the principal ancient, medieval, and modern languages of each. The names of branches and subbranches are in boldface. Not all languages or language stages are represented.



Map 1.1 Geographical distribution of the major Indo-European peoples around 500 BC

then Tocharian, split off from the family first, and that the remainder of PIE underwent further common development before the other branches emerged has found increasing support in recent years and is in many ways little different from Sturtevant's original claim. We will discuss these matters further in chapter 9.

The evidence that can confidently be used to evaluate these claims is unfortunately sparse. This is partly due to the fact that the earliest documentation of IE languages (from the mid-second millennium BC) is, by current estimates, still over two millennia later than the date of PIE itself. The branches therefore had considerable time to undergo changes before their first attestation, changes that could have obscured earlier developments that would help identify any subgroups. Also, some branches (such as Anatolian and Indo-Iranian) enter into recorded history much earlier than others (such as Baltic and Albanian, nearly three thousand years later), meaning they are not all at a comparable stage of development when we get our first glimpses of them.

In view of these difficulties, the branches of Indo-European are presented in this book as separate entities, following the traditional model. However, we will discuss the more prominent subgrouping hypotheses in later chapters as needed.

Indo-European and other language families

1.17. Indo-European is merely one of hundreds of known language families, and PIE was merely one of hundreds of languages (or more) spoken during the late Neolithic. PIE is therefore not the primeval ancestor of all languages now spoken, nor even of a majority of them (although there are more speakers of IE languages than of the languages of any other family). At the present time, we do not know whether any other language families share a common ancestor with PIE. The comparative method can give us solid results for a time depth of perhaps eight thousand years; most estimates of the age of PIE are in the neighborhood of six thousand years. (The age of PIE will be treated in detail in the next chapter.) Sporadic, irregular, and unpredictable linguistic changes such as analogy, paradigm leveling, borrowing, and semantic

change all make reconstruction more difficult the farther back in time one goes since they add “background noise”: borrowings become indistinguishable from native words, semantic changes make it difficult to reconstruct word meanings, and morphological analogies and leveling processes erase evidence of earlier inflectional patterns.

1.18. These problems have not prevented some adventurous souls from positing relationships between IE and other language families, or from coming up with classifications of huge numbers of language families into a handful of super-families. Chief among the former claims – the latter do not concern us – is the so-called *Nostratic* hypothesis (from Latin *nostrās* ‘one of us’, from *noster* ‘our’): this is the name of a putative language family that includes such groups as Indo-European, Uralic (Finnish, Estonian, Hungarian, etc.), Semitic (Arabic, Hebrew, Ethiopic, etc.), Dravidian (Tamil, Telugu, Kannaḍa, and other languages of southern India), and Kartvelian (Georgian and its relatives in the Caucasus). Supporters of the Nostratic hypothesis claim that Proto-Nostratic reconstructions help to elucidate the pre-history of PIE. But most Indo-Europeanists are not persuaded by it, and find that the Nostraticists’ methods lack the required rigor; at best the Nostratic hypothesis is premature.

This is not to say that the Nostratic hypothesis (or others like it) is necessarily wrong, or might not yield interesting results in the future if the methods are improved. One cannot disprove the claim that two languages or language families are related; but the more distant and obscure the relationship, the more vacuous the claim. Past a certain point, attempts at classification become uninteresting for scientific or historical purposes. The language relationships that have the most to tell us are those whose histories can be traced in some detail and with reasonable certainty. The comparative method seeks to explain why certain languages have systematic similarities too regular and too numerous to be due to chance. Indo-European does indeed share similarities with other language families, but to most eyes these similarities are too scattered and occasional to warrant application of the comparative method.

Conclusion

The limits of – and on – our knowledge

1.19. How complete is our picture of PIE? We know there are gaps in our knowledge that come not only from the inevitable loss and replacement of a percentage of words and grammatical forms over time, but also from the nature of our preserved texts. Both the representative genres and external features such as writing systems impose limits on what we can ascertain about the linguistic systems of both PIE and the ancient IE languages. As regards the first, our corpus of ancient IE texts is marked by a preponderance of poetic or literary works. These were often the ones that were deemed the most sacred or important, and were therefore carefully preserved. This means that specimens of ordinary language are rarer overall, which affects the nature and completeness of our picture of PIE in ways that we cannot always be aware of.

As regards the second point, the ancient (and of course also the modern) IE languages are written in a diverse assortment of writing systems, not all of which are

equally well-suited for representing the finer details of pronunciation. Alongside the very serviceable alphabets and clear spelling systems used, for example, by the ancient Greeks and Romans are the syllabaries that the Hittites and some other peoples used. Syllabaries, as will be discussed further in chapters 9 and 12, are poorly suited for unambiguously representing certain combinations of speech sounds, and do not always encode all the sounds present in a language to begin with (as is the case with Linear B, used to write Mycenaean Greek). This often puts the modern researcher at a considerable disadvantage.

1.20. While these hindrances are very real, they do not impede progress in the field outright. As of this writing, we approach the threshold of the third century of Indo-European studies (or, dating from Sir William Jones, we have long since crossed it), and our picture of PIE and the histories of its daughters has constantly undergone revision since the field's inception. We can rest assured that that picture will continue to grow and change. For this reason, the account of linguistic prehistory given in this book is not an immutable truth. Some of the most basic and long-held assumptions about PIE grammar have been challenged and even overturned in recent decades. In the 1950s and 1960s, Noam Chomsky of MIT founded the field of generative linguistics, which has revolutionized the study of language and made a whole host of new insights and theoretical apparatus available. The marriage of the traditional methods of comparison and philology with the findings of generative linguistics has in some cases led to the solution of long-intractable problems in Indo-European, and will without doubt yield great advances to come.

Since no documents in reconstructed PIE are preserved or can reasonably hope to be found, the structure of this hypothesized language will always be somewhat controversial. Of necessity, then, this book contains views on problems that still await final resolution; but most of the historical accounts presented here are subscribed to by a majority of researchers in the field, and care has been taken to flag any views for which this is not the case.

In spite of all the scholarly disagreements that enliven the pages of technical books and journals, all specialists would concur that enormous progress has been made since the earliest pioneering work in this field, with consensus having been reached on many substantial issues. The Proto-Indo-Europeans lived before the dawn of recorded human history, and it is a testament to the power of the comparative method that we know as much about them as we do.

The goals of Indo-European linguistics

1.21. Much of the preceding discussion has focused on the reconstruction of Proto-Indo-European, which is often imagined to be the field's main goal. Yet this is only one of the activities in which Indo-Europeanists are engaged. Reconstruction in many ways is important not as an end in itself, but because it provides a necessary means for elucidating the histories and properties of the attested daughter languages. This includes the modern languages, which – it should not be forgotten – are later and altered forms of the same ancient tongue that Latin and Greek are altered forms of. Regardless whether one is more interested in reconstruction of PIE or in developing historical linguistic theories to explain post-PIE linguistic developments, the exegesis of our ancient texts using philological principles is of fundamental importance, and much research in IE studies has always been devoted to that

endeavor. In no other way can we refine our understanding of the primary evidence from the languages upon which we base our comparisons, reconstructions, and historical accounts.

A further goal of many Indo-Europeanists is not a linguistic one: the reconstruction of PIE culture, and tracking the cultural developments of the descendant IE lineages. Much can be said about these things purely on the basis of comparative linguistic evidence. As a prelude to outlining the structure of the language of the Proto-Indo-Europeans, in the next chapter we will explore what we know about their culture, and follow some of the attempts to locate them in time and space.

For Further Reading

(The references are listed in the Bibliography, together with additional ones not commented on in this and the other “For Further Reading” sections.)

The classic treatment of the comparative method is Meillet 1925. A good introduction to historical linguistics for non-specialists is Sihler 2000; a bit more specialized is Campbell 1999. From the literature on the Nostratic hypothesis, a useful recent survey is Salmons and Joseph 1998; see therein the articles by Brent Vine and Lyle Campbell. A superb description of the evolution of historical linguistics and Indo-European studies into the beginning of the twentieth century is still Pedersen 1959, to which may also be added chapter 7 of Robins 1997.

As for general overviews of Indo-European linguistics, still the only complete treatments are the nearly century-old work of Brugmann 1897–1916 and the slightly later and smaller Hirt 1927–37, which in spite of their age have useful collections of data. Single-volume introductions to the field include Meillet 1937 (a celebrated classic), Beekes 1995 (readable but idiosyncratic), Szemerényi 1996 (quite idiosyncratic but filled with useful bibliographies after each section), and Meier-Brügger 2000 (whose translation [2003] is probably the best overview currently available in English and has an extensive bibliography, but is too difficult to use for the beginner). Bader 1997 and Ramat and Ramat 1998 are both collections of essays by different authors that treat the individual branches; many of the articles are excellent.

The most prominent scholarly journals devoted to Indo-European are generally published in Europe: *Bulletin de la Société de Linguistique de Paris*, Paris; *Historische Sprachforschung* (formerly *Zeitschrift für Vergleichende Sprachforschung* or *KZ* for short, since it was once *Kuhns Zeitschrift für Vergleichende Sprachforschung*), Göttingen; *Indogermanische Forschungen*, Berlin; *Journal of Indo-European Studies*, Washington; *Münchener Studien zur Sprachwissenschaft*, Munich; *Die Sprache*, Vienna. There is also a growing list of online resources. First is the TITUS project (short for *Thesaurus Indogermanischer Text- und Sprachmaterialien*, or *Thesaurus of Indo-European Text and Language Materials*) led by Jost Gippert at the University of Frankfurt (titus.uni-frankfurt.de), which will, if it is completed, have electronic versions of all the ancient and medieval Indo-European texts in searchable form; quite a few of them are already available, as well as other related information. The homepage of the Indo-European Institute at the Free University of Berlin (www.fu-berlin.de/indogermanistik) keeps up-to-date links to the major European and American research centers in Indo-European, and to announcements, bulletins, and many other sources of information.

For Review

Know the meaning or significance of the following:

comparative method	cognate	daughter language	Sir William
genetically related	regularity of sound	Neogrammarian	Jones
reconstruction	correspondences	hypothesis	Indo-Hittite
correspondence set	proto-language	philology	Nostratic

Exercises

- Memorize the names of all the branches of the IE family, and the names and filliations of all the extinct languages in figure 1.1.
- Below are given the names of the cardinal numerals 1–10 in nine languages. Based on this information only, do you think all the languages are genetically related? If not, which one(s) do you think is/are not related to the rest? Give evidence for your answer in a short paragraph. Ignore diacritic marks over and under the letters.

	A	B	C	D	E	F	G	H	I
1	heis	ipūn	yksi	yek	un	ains	mi	sas	kahi
2	dúō	lū	kaksi	do	dau	twai	erku	wu	lua
3	treis	čī	kolme	se	tri	threis	erek ^c	tre	kolu
4	téttares	čpū	neljä	cahār	pedwar	fidwör	čork ^c	štwar	hā
5	pénte	wuču	viisi	panj	pump	fimf	hing	pāñ	lima
6	héks	wušu	kuusi	šeš	chwech	saihs	vec ^c	šāk	ono
7	heptá	sete	seitsemän	haft	saith	sibun	evt ⁿ	špāt	hiku
8	októ	āstē	kahdeksan	hašt	wyth	ahtau	owt ^c	okāt	walu
9	ennéa	nū	yhdeksän	noh	naw	niun	inn	fiu	iwa
10	déka	lez	kymmenen	dah	deg	taihun	tasn	šāk	'umi

- It is not infrequently claimed that the tools developed for comparative reconstruction of Proto-Indo-European do not work outside the Indo-European family – as, for example, when working with the indigenous languages of Papua New Guinea, which have been stated to be very different from one another and spoken by a great variety of different ethnicities. Based on the discussion of the comparative method in this chapter, analyze this claim critically. What might be another explanation for the failure of the comparative method as applied to these Papuan languages?
- Remains of a previously unknown ancient language are unearthed in western Asia. How could one determine whether it is Indo-European?

2 Proto-Indo-European Culture and Archaeology

Introduction

2.1. In the previous chapter, we saw how the comparative method is used to reconstruct extinct languages, and in the next few chapters we will see specifically what it has accomplished in reconstructing the structure of Proto-Indo-European (PIE). But the comparative method has other applications, too. In its ability to reconstruct a prehistoric people's vocabulary, it opens up a valuable window onto their culture. A language does not exist apart from a people, and it always mirrors their culture to some extent. Furthermore, we can broaden the scope of comparison to include not only individual words but also their use in context, which reveals the semantic and cultural associations that attend different concepts. Thus comparative linguistic study allows us to reconstruct a *proto-culture* alongside the proto-language. (As with the term *proto-language*, there is nothing more "primitive" or "unformed" about a proto-culture; the term simply refers to a prehistoric culture which we know about by virtue of having reconstructed its language.)

Besides comparing linguistic forms, much effort has been devoted to the comparison of myths, laws, and all manner of social institutions. But if a set of related daughter cultures shares a particular myth, custom or the like, the attribution of that myth or custom to the proto-culture is more secure if it is accompanied also by a *linguistic* equation. This is not to say that linguistic comparison is indispensable for cultural reconstruction; but without it, it can be harder to dismiss the possibility of independent innovation or borrowing on the part of the daughter traditions. The comparative method remains the most powerful tool available for bringing the non-material culture of prehistoric peoples to light.

Below we will sketch what has been learned about PIE culture and society from the reconstructed vocabulary of PIE and the cognate cultural traditions of the daughter branches. We will follow this with a discussion of the great (and notorious) question of the location of the PIE homeland and the allied question of the date of the breakup of PIE. Because space is limited, and because we have not yet introduced the notational conventions used for spelling reconstructed PIE forms, specific reconstructions and lists of descendant forms will be almost entirely eschewed. Note also that in most cases, comparanda for any given cultural institution, myth, etc. are only given selectively; where an illustrative example of a particular item is given from only one or two branches of the family, it should not be assumed that that item is represented only in those branches.

2.2. Before embarking on these discussions, it should be noted that an important question is begged by such endeavors as reconstructing PIE culture, locating the PIE homeland, and dating the end of PIE linguistic unity. These pursuits all assume that there was at one time a fully homogeneous and reifiable PIE language and culture that suddenly ceased to exist as such. In fact, we know quite well that this is not true of any speech community or culture, and that linguistic “breakups” are gradual processes. Science often finds it necessary, however, to distance itself from the messiness of the real world and to deal in idealizations; and that is what must be done here. The true heterogeneity of the PIE speech community is not something we can possibly recover; but what we can recover is a picture of what PIE speakers had in common, both linguistically and culturally. It is not the business of comparative linguistics to reconstruct a panoply of individual variation or even to worry about it, for that would strip the whole notion of a “common ancestor” of any meaning. The temporal side of all of this is that we cannot hope to know, except maybe in a few important cases to be discussed below, which reconstructed words belong to which chronological layer of the proto-language. Suffice it to say that the words discussed in this chapter are ascribed to PIE by a majority of specialists, rather than just to some later dialect area that postdated the common period (that is, the period of PIE linguistic unity).

Society

Social stratification and organization

CLASSES OF SOCIETY

2.3. It is universally agreed that PIE society was hierarchical. First, there was a general distinction between free persons and slaves; the latter, as in many non-IE societies, were typically captives taken in war or debtors unable to repay a debt. (Words meaning ‘man, warrior’ came secondarily to mean ‘slave’ in some traditions.) The free segment of society was further subdivided into an elite class of kings, warriors, and priests (and probably poets; cf. §§2.37–38) on the one hand, and into a class of common people on the other. These distinctions had legal repercussions: in Old Irish law, for example, an injury to a person of high rank demanded a greater penalty than the same injury to one of low rank. Additionally, men outranked women; the society was patriarchal, patrilineal, and patrilocal (with brides going to live with the family of their husbands, on which more in §2.6 below).

2.4. One of the most influential structural approaches to analyzing PIE society is that propounded by the twentieth-century French Indo-Europeanist Georges Dumézil. In his view, PIE society, especially in the form of its free males, was divided into three basic aspects or “functions.” The first function encompassed both sovereignty and religion, and was embodied in priests and kings that kept religious and legal order. The second function was that of martial force and was represented by the warrior class. The third function was that of fertility, embodied in pastoralists and in other producers of goods (artisans, for example).

Some of the early IE societies are indeed divided up along these lines, most famously in ancient India and Iran. The traditional caste system in India, which divides society into priestly, warrior, and herder-cultivator classes (plus a fourth into which were

originally relegated the subjugated non-Indic peoples), is already mentioned in the Rig Veda, the oldest Sanskrit (Old Indic) text, and in Dumézil's view is a direct continuation of the three functions. But elsewhere in the IE world such clear three-fold divisions are difficult to come by. The ancient Celtic society of the Gauls, as described by Julius Caesar, consisted of priests, knights, and a nearly slavlike commonfolk; and descriptions of certain Ionic and Doric Greek tribal divisions agree with the general model, but seem to have been rather marginal.

As Dumézil developed his theory, he grew to envision these functions less as actual divisions of society, and more as composing a (not always clearly defined) kind of cognitive framework, an "ideology" (*idéologie*), which he claimed was reflected in the structure of the pantheons of different IE religions, in myths, in religious practices, and in other cultural arenas across the Indo-European-speaking world. We will have occasion to examine and evaluate this application of his framework in §§2.35–36 below.

KINSHIP AND THE FAMILY

2.5. The kinship system of the Indo-Europeans is fairly well understood. The PIE words for father, mother, brother, sister, son, and daughter have descendants in almost every branch; also reconstructible are words for grandfather, mother's brother, and nephew and niece. The terms for 'nephew' and 'niece' also meant 'grandson' and 'granddaughter'; all of these are agnates at least a generation younger and two relations removed from the person of reference. Words for several more complex relationships can also be reconstructed, including daughter-in-law, husband's father, husband's mother, husband's brother, and brother's wife. A few others are a bit more uncertain: Indic and Slavic have cognate words for wife's brother, though a third cognate, in Armenian, means 'son-in-law'; and two languages, Old Norse and Greek, apparently preserve an inherited term for wife's sister's husband.

As the evidence above shows, more kinship terms for males or relatives of male kin can be reconstructed than for females. Anthropologists have classified the kinship systems of the world into several basic types; the PIE system fits none of these exactly (and they are ideal constructs anyway), but the closest match is the one known as the Omaha system. This system is found in patrilineal exogamous societies, that is, those where descent is reckoned through the father's line and spouses are taken from outside the kin group.

2.6. No single term for 'marriage' can be reconstructed; different legal kinds of marriage were recognized, including marriage by abduction. Specific procedures had to be followed for each of them. In the daughter languages, 'to marry' (a woman) is usually expressed by a verb meaning 'lead away' or 'take' (as Latin *uxorem ducere* 'lead a wife, marry'), and this can be confidently projected back onto the proto-language; the relevant roots are used also of cattle or water, and their use here indicates exogamous and virilocal marriage where the bride was 'taken' or 'led' from her father's family to her husband's. (For this reason, the PIE word for daughter-in-law came to mean 'bride' in Armenian and Albanian; from the husband's family's point of view, the daughter-in-law was the new bride in the family.) In PIE society the husband's family had to pay bridewealth (also called bride price), the word for which has descendants in several branches.

2.7. Fosterage was surely practiced in PIE times, just as it was in many of the daughter cultures. In several cases, the relationship to one's foster-father was closer

than to one's natural father. Thus in Old Irish the inherited words for 'mother' and 'father' (*máthair*, *athair*) refer to one's biological mother and father, whereas the more affectionate baby-talk words *muimne* and *aite* refer to one's foster-mother and -father. Foster-parents were chosen preferentially from the mother's kin; the maternal uncle was particularly common in the role of foster-father.

SOCIAL UNITS

2.8. Aside from having class divisions, PIE society also consisted of small units organized into larger ones. Here, though, there is no agreement among scholars on the specifics, since about a half-dozen words for social units from the household on up can be reconstructed, but their precise meanings are uncertain. None seems to have referred to anything more extensive than the clan, except perhaps the word **teutā-*, meaning 'people, tribe', which has descendants in Italic, Celtic, Germanic, and Baltic, and is probably also found in personal names in Thracian, Illyrian, and Messapic. But as this word is confined to European languages, its status as a PIE inheritance is uncertain. Nonetheless, it has recently been proposed that the **teutā-* was the central unit of PIE social organization, with a division between those outside and those inside the tribe. According to this theory, proposed by the Celticist Kim McCone, certain adolescent males would join a warrior-band (or, as it is frequently called, a *Männerbund*, the German term) that engaged in various acts of violence (including raiding and pillaging), for which they were identified symbolically with wolves. Under this view, society was fundamentally structured around the organization of warfare.

2.9. We can reconstruct words for leaders of at least three ranks, up to what is usually translated as 'king' (the source of Latin *rēx* and Gaulish *rīx*), who was at the head of the **teutā* in those languages which knew the term. No self-designation of the Proto-Indo-Europeans has survived (there may have been no special term); it was formerly thought that the Indo-Iranian tribal self-designation, *ārya-* (Aryan), was the continuation of such a term in PIE, but this theory is no longer generally accepted. We will discuss this term in more detail in §10.28.

Economics and reciprocity

TYPES OF PROPERTY

2.10. The older IE languages typically distinguish between movable and immovable wealth, and in the former category between two-footed and four-footed chattels, with humans being the two-footed kind. We can reconstruct a general word for 'movable wealth', which in several languages became specifically the word for live-stock, movable wealth *par excellence* in a pastoral society (e.g., Latin *pecū*, Old High German *fihu*). But it retained its general meaning of 'wealth' or 'money' ultimately in English *fee*.

Property was probably divided into hierarchical categories that had legal relevance. Ancient Roman law classifies small livestock, large livestock, men, and rights to land in a separate category from other types of property; and within this category these types of property form a hierarchy where small livestock is at the bottom and land is at the top. This exact same hierarchy occurs in Indo-Iranian legal tradition, and is likely inherited.

EXCHANGE AND RECIPROCITY

2.11. Various roots having to do with transaction, buying and selling, payment, and recompense have been reconstructed. They attest to a well-developed economic exchange system, one of the aspects of IE society that revolved around reciprocity. A gift always entailed a counter-gift, an exchange always involved a mutual transaction; this simple principle was manifest in the meanings of the central terms of exchange, which – it has been argued – did not mean simply ‘give’ or ‘take’ but referred to the whole act involving both parties of the exchange. For this reason, such roots have descendants that refer to one side of the exchange in one set of daughter languages and to the other side in other daughters: Greek *németai* ‘allots’ is cognate with German *nehmen* ‘take’; Tocharian B *ai-* ‘give’ is cognate with Greek *ainumai* ‘I take’; and so forth.

Reciprocity was manifest in virtually every corner of PIE society – in the relationship between the two parties in a contractual agreement, between guest and host (see the next paragraph), poet and patron (§2.38 below), and gods and humans (§2.37 below). These may seem like fundamentally different interactions, but not from the PIE point of view: each party to these relationships was mutually bound to the other, and the relationship was cemented (and only made possible) by trust. Derivatives of the PIE root for ‘trust’ are widespread, and include words referring to that concept (such as Latin *fidēs*) as well as to particular types of mutual agreements bound by trust (everything from Latin *foedus* ‘treaty’ to Albanian *besë* ‘truce in a blood feud’). These trust-based institutions transcended the boundaries between economics, law, and religion.

2.12. The institution of hospitality, the guest–host relationship, is a case in point. As far as we can tell, PIE did not have words distinguishing ‘guest’ from ‘host’; rather, there was a single term meaning something like ‘a stranger with whom one has reciprocal duties of hospitality’. The giving and receiving of hospitality was accompanied by a set of ritual actions, including gift-giving, that indebted the guest to show hospitality to his host at any time in the future. The obligation was even heritable, making guest-friendship practically a kind of kinship. A famous passage in the *Iliad* describes an encounter between the Lycian warrior Glaukos (fighting for the Trojans) and the Greek warrior Diomedes that nicely illustrates this principle. In the encounter, Glaukos and Diomedes tell each other the story of their lineages, whereupon they discover that Glaukos’s grandfather had once been a guest at the house of Diomedes’s grandfather. Upon discovering this, the two decide not to fight each other, and instead exchange armor and renew the vow of guest-friendship inherited from their grandfathers. The exchange of armor repays the old debt: Glaukos’s armor is much more valuable than Diomedes’s. (The narrator of the tale, interestingly, seems not to understand the proceedings and claims that Glaukos’s wits were added.)

Violations of the guest–host obligation were illegal, immoral, and unholy. In Irish law, refusing to give hospitality was a crime that demanded payment of the offended person’s full honor-price, the same penalty exacted for serious injury and murder. The killing of a guest in IE societies was greeted with singular revulsion, and is the fertile subject of many legends. In the *Odyssey*, what made the killing and eating of some of Odysseus’s men by the Cyclops so revolting was that they were the Cyclops’s guests. Hospitality could be abused too; the Trojan prince Paris, by abducting Helen, the wife of his host Menelaos, was perhaps the ultimate bad guest, and the *Odyssey* spends considerable time developing the motif of the suitors,

“anti-guests” who camped out in Odysseus’s home in his absence while suing for the hand of his wife, Penelope.

Law

2.13. The study of legal vocabulary is important for IE linguistics because the archaic nature of traditional legal phraseology preserves old forms and meanings of words that are often not preserved elsewhere. As with religious formulations, legal formulations must be uttered precisely the same way each time to be binding; the Roman jurist Gaius (fl. second century AD) gives an example of a lengthy legal formula in its legally binding version and in a minimally different version that, he tells us, is legally worthless.

Relatively little work has been done on the comparative reconstruction of PIE law and legal vocabulary. This is not for want of material, which is abundant and includes (among other things) the Hittite Law Code, the law of Manu in Vedic India (the *Mānavadharmasāstra*), the Gortynian Code from Crete, the Laws of the Twelve Tables of ancient Rome, numerous Old Irish legal texts, northern Albanian customary law (the Code of Lekë Dukagjini), and various medieval Germanic and Slavic law codes. Rather, there are certain problems inherent in the texts themselves. One is that outside influence must always be reckoned with; for example, the Hittite Laws contain elements that are common to other (and non-IE) ancient Near Eastern societies. Another is that laws that have been codified and written down represent, at least in part, legal reform rather than untouched ancient practice.

2.14. These problems can be easily overemphasized, however. Careful comparative linguistic study of legal phraseology in cognate traditions can uncover, and has uncovered, inherited legal vocabulary and idioms – and with it, PIE legal practice. In Hittite and Roman law, restitution for damages done by one’s son or slave to another party was achieved by the father or master paying for the damages himself or surrendering the perpetrator to the offended party. This act of compensation is expressed in Latin by the verb *sarcīre* and in Hittite by its cognate, *šarnik-* (the *-ni-* is not part of the root). Since the two traditions agree precisely in both legal and linguistic content, it is safe to assume that the PIE root **sark-* that underlay *sarcīre* and *šarnik-* had a technical legal usage in referring to this particular type of recompense.

2.15. Future studies will surely uncover many more technical IE legal terms of this kind; in the meantime, there is not much that we can say about this corner of the PIE lexicon. We are not even sure what the general term or terms for ‘law’ were. A word probably meaning ‘law’ or ‘religious law’, originally in the sense of ‘legal or ritual statement that must be pronounced’ or the like, has been reconstructed on the basis of Indo-Iranian and Italic; it is the source of Latin *ius* ‘law’ (whence English *justice*). The verb meaning ‘place, put’ apparently had legal overtones; it furnishes such derivatives as *dhāma* ‘law’ in Vedic Sanskrit and *thémis* ‘law’ in Greek, words that were applied especially to laws at the level of the family or household. The notion was that of something ‘placed’ or established. English *law*, a borrowing from Old Norse, comes from a root meaning ‘lay’ and is therefore ‘that which is laid down’; it is possible that Latin *lex* comes from the same root, though this is debated. We cannot reconstruct a word for the central concept of the ‘oath’, the swearing of which was both a religious and a legal act (as it is today when one swears by the Bible to tell the truth in a court of law); each branch has a different term.

2.16. In PIE society, there was no public enforcement of justice. In order for contractual obligations to be met, private individuals probably acted as sureties (that is, they pledged to be responsible for payments of debts incurred by someone else in case the latter defaulted). The fact that there were no higher officials that enforced justice meant that individuals had to take matters into their own hands sometimes; in Irish law and the Roman Laws of the Twelve Tables, one could formally bar someone from access to their property to compel payment. PIE society probably knew no formal court as we know it today, but suits could be brought by one party against another, and cases were argued before judges (perhaps kings) that featured witnesses. Irish and Gothic preserve what might be an inherited term for 'witness' that is a derivative of the verbal root meaning 'see' or 'know'. Italic has famously innovated a term meaning 'third person standing by' (*testis*, from earlier **tri-stis*), for which there is a near-equivalent in Hittite: a compound verb meaning 'stand over' (*šēr ar-*) had an extended technical meaning 'bear witness'.

Religion, Ritual, and Myth

Indo-European deities

2.17. All the older IE religions are polytheistic, as was that of the Proto-Indo-Europeans. Nothing like a complete picture of PIE religious beliefs and practices is possible; in what follows, we can only give a sampling of the major divine figures, myths, and a few elements of religious ritual. On the whole, few divine names can be confidently reconstructed. Most of the familiar Greek and Roman gods, for example, have names of unknown etymology, and some (like Aphrodite) are known to be of Semitic provenance. Others, like Venus and the Germanic god Woden, have names that derive from Indo-European roots, but there are no deities in other branches with cognate names. Clearly the daughter traditions have undergone considerable change and evolution.

2.18. Some idea of how the Proto-Indo-Europeans conceived of their relationship to the gods can be seen in the etymology of their term for 'human being', whose descendants include Latin *homō* and Old English *guma* (the latter preserved in altered form in the compound *bride-groom*): the PIE form was derived from the word for 'earth' or 'land', attesting to a conception of humans as 'earthlings' as contrasted with the divine residents of the heavens. Another paired contrast is evident in the widespread use of the word for 'mortal' as a synonym for 'human', as opposed to the immortal gods.

Given that gods were in the first instance celestial beings in the IE view of the cosmos, it is not surprising that the most securely reconstructible members of the PIE pantheon had to do with the sky and meteorological phenomena; they were also mostly male (but see below). The general word for 'god' is a derivative of a root meaning 'shine', as of the bright sky; its descendants include such words as Vedic *devás*, Latin *deus*, Old Irish *día*, and Lithuanian *diēvas* (but not Greek *theós* 'god', which is from a different root).

2.19. The same root for 'shine' furnished the name of the head of the PIE pantheon, a god called **Father Sky**, whose name is securely reconstructible from the exact equation of Vedic Sanskrit *dyāus pítar* '(o) Father Sky', Greek *Zeū páter* '(o) Father Zeus', and Latin *Iū-piter* 'Jupiter' (literally 'father Jove', also originally a

vocative or form of direct address like the previous two). Compare also Luvian *tatiš Tiwaz* 'father Tiwaz', where the deity has been transformed into a sun-god, and Old Irish *In Dagdae oll-atbir* 'the Good God, super-father'. In Germanic, the head god became the god of war: Old Norse *Týr*, Old English *Tig*, whence our *Tues-day*. The appearance of the word for 'father' as part of the IE Sky-god's title probably referred to his hierarchical position at the head of the pantheon, and not necessarily to any role as progenitor. It would then be like the same use of Latin *pater* 'father' in the ancient phrase *pater familiās* 'head of the household'.

2.20. Alongside Father Sky was another male deity, the Sun. His daily course across the sky is envisioned in many IE traditions as a horse-drawn chariot ride; though probably inherited, this motif cannot be more ancient than the invention of wheeled vehicles, which were a late addition to PIE culture (see §2.58 below). In the western IE branches, the Sun and associated deities rested on an island in the western sea after their daily journey, an island frequently described as having an apple-orchard and sometimes associated with the realm of the dead. Whether this is an inherited motif is uncertain.

2.21. The Indo-Europeans had a god of thunder and lightning, probably represented as holding a hammer or similar weapon; this is how the Baltic thunder-god Perkunas and the Old Norse god Thor are depicted (the name of the latter's hammer, *Mjöllnir*, is cognate with words in Celtic and Balto-Slavic for 'lightning'), and also in some representations the Anatolian Stormgod. Thunder and lightning have both destructive and regenerative associations; a lightning bolt can cleave stone and tree, but is accompanied by fructifying rain. This gives rise to the folk-motif of the lightning bolt that impregnates rocks and trees (especially the oak), and explains the strong associations between the Balto-Slavic god of thunder (Lithuanian *Perkūnas*, Old Russian *Perunū*) and the oak. It is interesting to note in this connection that the PIE word for 'stone' secondarily refers to 'heaven' in Indo-Iranian and Germanic; while we are not entirely certain of the underlying association, it may rest on a conception of the heaven as a stony vault, from which fragments might fall in the form of meteorites; or it may be connected with the stony missiles thought to be hurled by the god of thunder.

2.22. Alongside these male sky and weather gods, we know of at least one goddess in PIE mythology, the Dawn, whose Indo-European name becomes *Uṣās* in Vedic myth, *Ēós* in Greek, and *Aurōra* in Latin. In three traditions (Indic, Greek, and Baltic) she is also called the 'daughter of heaven', perhaps an inherited epithet; and in these three branches plus a fourth (Italic) there is a story of the reluctant dawn-goddess who is chased or beaten from the scene for tarrying. The Indo-Europeans oriented themselves by facing east, toward the dawn, as shown by the fact that 'south' in PIE was expressed by the word for 'right'. ('East' itself was expressed by the word for 'dawn' or a derivative of it, and similarly 'west' was expressed by the word for 'evening'.)

Other goddesses have been proposed for PIE as well, but they are less certain. Most daughter branches have a 'Mother Earth', a figure ubiquitous around the world and not specifically Indo-European, but within the IE family itself there is no reason why her name could not be inherited.

2.23. An important pair of figures in IE myth are the divine twins (their names cannot be recovered), whose most familiar representatives are Castor and Pollux in Greco-Roman myth (the Dioskouroi, 'sons of Zeus'). They are also continued by the *Nāsatyas* or *Aśvins* in Vedic India and as the *Dieva dēli* or 'sons of heaven' in Latvian folklore. From the considerable mass of often quite varied legends that surround these

figures, it appears that the PIE divine twins were offspring of Sky, were youthful, and were connected with (or even took the form of) horses, especially the horses that drew the chariot of the Sun. More distant echoes of the horse-twins are arguably found in the Irish legend of Macha, who gave birth to twins after winning a footrace against horses, and in the Germanic figures Hengist and Horsa, legendary or semi-legendary leaders of the Anglo-Saxon invaders of England and founders of the kingdom of Kent (their names mean 'stallion' and 'horseman', respectively).

2.24. The opposing elements of fire and water are intimately associated in the divine Indo-Iranian figure called the 'grandson (or nephew) of the waters' (Vedic *Apām Nápāt*, Avestan *Apam Napā*), depicted as a fiery god residing in the water, giving off light, and needing to be propitiated for the proper use of bodies of water. An Irish mythical figure, Nechtan, is the subject of a myth with similar elements; etymologically, his name can be reconstructed as **neptonos*, formed from **nept-*, one of the stems of the PIE word for 'grandson' or 'nephew' that gives *napāt* in Indo-Iranian. This same stem **nept-* recurs in the name of the Roman god of waters, *Neptūnus* (Neptune); in Roman myth there is no longer any overt connection between him and fire, although there are tantalizing traces of an old Neptune cult with elements recalling those found in the cults of *Apām Nápāt* and Nechtan. The Armenian tale known as the Birth of Vahagn (see §16.42) is analogous to these in featuring fire born of a water-dwelling plant.

Ritual and cultic practice

FIRE-WORSHIP

2.25. An interesting fact of the reconstructed PIE lexicon is that 'fire' and 'water' could each be expressed by different terms, one of animate gender and one of inanimate gender; this has been taken to reflect two conceptions of fire and water, as animate beings and as substances. The most dramatic reflection of the former is in the deification of fire that is seen in various IE traditions. The evidence for fire-worship as part of PIE cultic practice is scanty, but compelling in its details. Vedic India worshiped the fire-god Agni (literally 'Fire'), and the ancient Iranian Zoroastrians were famous fire-worshippers, though their fire-god had a different name, *Ātar* (also literally 'Fire'). The Romans divinized the domestic hearth and its fire in the form of the goddess Vesta. Though the names of these deities are all different, the temple of Vesta that housed her sacred fire, uniquely among Roman temples, was circular rather than square; and the domestic fire in ancient India was accorded a round altar, rather than the square one used for public worship.

KINGSHIP RITUAL AND THE HORSE SACRIFICE

2.26. Indic, Roman, and Irish traditions, and indirectly also Anatolian, attest to an important ritual held to consecrate kingship whose central act was horse sacrifice. We know far more details about the Indic ritual, called the *āsvamedha*, than the others; its core elements were the sacrifice of a stallion (specifically, one that excels on the right-hand side of the yoke), ritual copulation with the dead stallion by the queen, and the cutting up and distributing of the horse's parts. We have traces of an ancient Roman ritual called the October Equus, which involved the sacrifice (to Mars) of the right-hand horse of the victorious team in a chariot race, the cutting

off of its head to be fought over by two groups of people, and the affixing of its tail to the wall of the Regia (the ancient royal palace from when Rome was ruled by kings). In the twelfth century, a Welshman named Giraldus Cambrensis described in his *Topography of Ireland* a ritual among the Irish that involved the copulation of a king with a mare that is then killed and boiled and cut into parts, which are subsequently distributed to everybody to eat.

On the basis of these comparanda, one can conjecture that the Proto-Indo-Europeans had a ritual for the renewal of kingship involving the ritual copulation of a king or queen with a horse, which was then sacrificed and cut up for distribution to the other participants in the ritual. Some details may be of a later date, such as the Indic and Roman specification that the horse excel on the right side of the yoke: paired draft-horses do not appear in the archaeological record before the mid-third millennium BC – a date that (as we will discuss in greater detail later) is well after the breakup of PIE. In Anatolia, we have some traces of ritual royal copulation, but without horses; interestingly, though, in Hittite law, copulation with animals was a punishable offense except copulation with horses or mules.

THE AFTERLIFE

2.27. Several daughter traditions believed that the soul journeyed after death across a body of water to an afterlife. The journey undertaken could be arduous, and required prayers and offerings of food on the part of the soul's living kin, at least for a period of time. (It also required burying various goods along with the deceased that would be needed on the journey; see §2.65 below.) A Hittite ritual calls for pouring honey and oil onto the ground to "smooth out" the path for the soul. The journey on land culminated in reaching a body of water across which the soul had to be ferried, probably by an old man; the Greek myth of Charon and the river Styx is the most familiar descendant of this, but comparable myths are found in Celtic, Old Norse, and – with some modification – Indic and Slavic. There is no particular agreement across the different daughter traditions on what the underworld was like (according to one theory, it was originally conceived as a meadow, but this is disputed). In Greek, Germanic, and Celtic myth, a dog guards the entrance to the underworld (the 'hellhound'), and dogs are choosers of the dead in Indic and Celtic. In several traditions, underworld bodies of water are associated with memory, either taking it away (as the river Lethe in the Greek underworld) or imparting great wisdom (as the wellspring of Mímir in Old Norse myth).

MAGIC

2.28. We know comparatively little about magic in PIE times, although there is no doubt that it was practiced. Several branches attest the use of magical charms, spells, and curses. The Hittites used sympathetic magic involving the ritual manipulation of dolls and other objects or substances representing various evils; of central importance was the utterance of a spell, typically consisting of an extended simile (such as, "Just as this wax melts, and just as the mutton fat dissolves, let whoever breaks these oaths melt like wax [and] dissolve like mutton fat," from a text known as the Soldier's Oath). The Greeks, Italic peoples, and ancient Gauls left behind many prayer and curse tablets; the practice of writing curse tablets probably diffused from the ancient Near East, but the verbal artistry found in some of these spells has an

Indo-European flavor and may continue an inherited tradition. Some charms, such as certain ones against worms, are woven out of the same verbal fabric used in the telling of the dragon-slaying myth described below.

The use of spells and incantations was one of three categories of medical treatment in the ancient IE world, the others being the use of a knife or surgical instrument and the use of herbs or drugs. Texts in Vedic, Greek, and Celtic agree on this threefold division of medicine, and the use of incantations, according to the Vedic poet, is the best – a testament to the power of the word, on which more presently.

Myths

2.29. Certain aspects of religion are remarkably resistant to change. Religious formulae used in ritual, like legal formulae, must be worded just right to have the desired effect, and fixed religious phrases usually preserve archaic language. Also, the basic wording of myth narration is often exceptionally stable, even in the face of significant changes to or substitutions in the characters and events portrayed. A number of recent studies have shown that the specific words used in telling the kernel of a mythic tale are part and parcel of the myth itself. These words constitute the *basic formula*, the verbal vehicle encapsulating the myth. Thus, when two Indo-European cultures share not only a particular story but also particular formulaic words and expressions in telling that story, it can be shown that the story is inherited. A prominent example is the dragon-slaying myth.

THE DRAGON-SLAYING MYTH

2.30. Dragon-slaying myths are told the world over; thanks to research by the American Indo-Europeanist Calvert Watkins, the verbal and cultural elements that are specific to the IE version have now been detailed. The IE myth is directly continued, for example, in the Vedic Indic story of the god Indra (the head of the ancient Hindu pantheon) smiting the serpent Vṛtra to free the waters that the latter has trapped in his mountain lair. The story is simple on its face, but has deep significance: the waters are necessary for the health of the community; by hoarding them, the serpent has upset the natural order whereby wealth and nourishment are allowed to circulate, and Indra must thus do battle to restore order. The serpent as hoarder finds a close analogue in the well-known portrayal of dragons in Germanic legend (and in Tolkienesque derivatives thereof) as hoarders of treasure; their treasure-hoarding upsets the societal order by keeping wealth from circulating. In longer versions of the Hindu myth, Indra is in fact first defeated by the serpent; he must then get help from other deities who provide him with the intoxicating drink called soma to give him strength. This expansion is also an inherited motif: in the cognate Hittite legend, the storm-god Tarḫunnaš is at first defeated by the serpent, and only succeeds the second time around after drinking an intoxicant.

2.31. In PIE, this myth was encapsulated in the alliterative formulaic phrase **(e-)gʷhen-t ogʷhim* '(he) killed the serpent'. (The sound *gʷh* will be explained in the next chapter.) The root **gʷhen-* was one of several reconstructible roots for 'smite, kill', but a close study of the use of its descendants shows that it was reserved for acts of killing that involve a monstrous adversary, or acts of killing that are themselves monstrous and upset the natural order of things. (It was, in fact, quite comparable in usage to Modern English *slay*.) The formulaic language of the

dragon-slaying myth could thus be extended to a number of heroic exploits that did not involve dragons or serpents *per se*. In Greek literature, the same words are used of heroic slayings such as Bellerophon's of the Chimaera, and of slayings that upset the societal order such as Herakles's murder of his guest Iphitos, and Clytemnestra's slaying of her husband Agamemnon (where she in fact is overtly compared to a snake). Furthermore, Watkins was able to show that the formula was connected with another one that translates as 'overcome death' (compare §2.37 below), which has a whole host of other associations. Thus the words used as a vehicle for the serpent-slaying myth encapsulate not only that myth, but also a whole complex of cultural notions pertaining to the slaying of (or by) a monstrous opponent, the struggle of order against chaos, and rebirth.

CREATION AND FOUNDATION MYTHS

2.32. At least three traditions – Indic, Italic, and Germanic – have interrelated creation and foundation myths that involve the sacrifice of a primeval being named 'Twin' by a primeval man, and the carving up of a primeval man into the parts that make up the physical or social world. Norse myth tells of Ymir ('twin'), whose carcass was carved up by the gods to create the world. The motif of creating the world from the body of a primordial figure is quasi-universal, but the fact that the figure is named 'twin' is not. In Indic mythology, the primeval twin, Yama, was the first man to die, and his brother, Manu, was the founder of religious law. Manu means 'man', and another Germanic myth, reported by the Roman historian Tacitus, tells of the creation of three ancient Germanic tribes that represented the three classes of society by Mannus ('man') and his father Tuisto ('twin'). Resembling a combination of the Ymir and Mannus myths are two legends from ancient India and Iran. The first is the story of the creation of the four castes of Hindu society from the body parts of the primeval man Puruṣa (Puruṣa): his upper body parts became the upper castes, and his lower parts became the commoners (see the excerpt in §10.51). The second is the Iranian myth of Yama Xšaēta, who, as a consequence of sinning, had his triple halo taken away and distributed to the heads of the social classes, and was later cut in two by his brother.

The legends of the founding of cities or the origins of a people are often based on cosmogonic myths; if we turn to ancient Italy, we encounter another analogue, the myth of the founding of Rome by Romulus and his twin brother Remus and the latter's murder at the hands of the former (interpreted by some modern scholars as a primordial sacrifice that was necessary for the act of societal creation).

THEFT OF FIRE

2.33. PIE mythology evidently had a myth of the theft of fire. In Greek mythology, the titan Prometheus stole fire from the gods and gave it to the humans that he had just created. We usually read that the titan's name meant 'forethought', but that is simply a folk etymology on the part of the Greeks, who had long before forgotten the true meaning of his name. *Promēthēús* originally meant 'the one who steals'; it has an exact cognate, including the prefix, in the Vedic verb *pra math-*, which means 'to steal' and is used in the Vedic myth of the theft of fire. We cannot reconstruct the details of the PIE myth, but can assume that this particular compound verb was at the core of telling it.

ANIMALS IN PIE MYTH AND FOLKLORE

2.34. Several animals had mythological and folkloric associations. Goats draw the chariot of the god Thor in Norse mythology and the chariot of the Indic god Pushan, and they are associated with the Baltic god Perkunas. The wolf's name underwent *taboo deformation* (as we would say *gosh* for *God*) in several branches, suggesting it was feared; we also know it was associated in PIE culture with outlaws (cp. §2.8 above). In Hittite law, calling a person a wolf was a speech-act that legally branded the person an outcast. The wolf was also a symbol of death; seeing a wolf was a metaphor for being struck dumb, itself metaphorical for losing vitality and dying. (This is found in non-IE folk beliefs in Europe, too.) Probably owing to a hunters' superstition whereby uttering the name of one's quarry was forbidden for fear the animal might hear his name and make himself scarce, the bear's name was taboo in the northern European branches and was replaced by circumlocutions like 'the brown one' (Germanic), 'honey-eater' (Slavic), 'licker' (of honey, Baltic; the inherited word may survive in a term meaning 'bear's den'), 'honey-desirer' or 'good calf' (Irish; the inherited word survives as the personal name *Art*). Several daughter branches preserve a legend, perhaps inherited, of a mythical crane that devours an enemy people. The crow and raven were associated with prophetic knowledge in IE legend; both the Celtic god Lug and the Norse god Odin had two ravens that supplied them with information.

Dumézilian trifunctionalism and the interpretation of PIE religion

2.35. As discussed earlier, the trifunctional ideology proposed by Georges Dumézil (cf. §2.4 above) is said to be reflected in the structure of many aspects of IE religion. In early Roman religion, for example, a central trio of gods was formed by Jupiter the sovereign god (first function), Mars the god of war (second function), and Quirinus the patron of the common people (third function). In the second-millennium-BC Mitanni documents (see §§10.21ff.), which contain the first attested words in an Indic language, the names of the gods invoked at the signing of a treaty are Mitra and Varuna (sovereign and priestly first function), Indra (the warrior god, second function), and the divine twins the Nāsatyas (the third function). An Old Persian inscription of Darius the Great contains a prayer asking for protection from enemy onslaught (second function), poor crops (third function), and the Lie, the evil antithesis of religious Truth in the Persian Zoroastrian religion (first function). Similarly, an archaic Roman prayer (given in §13.53) contains an entreaty for warding off diseases (first function; medicine was part of the religious realm, since diseases were treated by spells, prayers, and the like), devastation caused by war (second function), and devastation caused by nature (third function).

2.36. The recognition of recurring structural similarities across such disparate material is arguably Dumézil's most notable achievement, and the tripartite ideology that he used to explain it has become standard doctrine among many specialists in comparative IE myth and culture. But there are good reasons to be cautious with it. In the first place, the ideology does not match very well much of the material that scholars try to apply it to. The deviations from expected trifunctionality have been explained away by the ad hoc postulation of various historical distorting influences. This introduces considerable interpretive flexibility that robs the theory of methodological rigor, and there is also a dangerous circularity in applying a theory to a set of

myths and analyzing those myths in such a way as to make them fit the theory. Most of the divinities in the various daughter cultures are complex figures that have facets belonging to two or even all three of Dumézil's functions; it is unclear how much understanding, or interest, is gained by reducing such rich cultural material to a framework that is ultimately rather bare. Even if the trifunctionalism of PIE is real, it may turn out to be a cognitive quasi-universal, as it has been documented for some non-IE cultures as well. It therefore remains to be seen how the recurring structural themes that Dumézil observed are best interpreted.

Poetics

Poets, patrons, and fame

2.37. The Indo-European poet was the society's highest-paid professional, specially trained in the art of the word. Not only was he a repository and transmitter of inherited cultural knowledge, but was also entrusted with singing the praises of heroes, kings, and the gods. Composing hymns in praise of the gods ensured that the gods would in turn bestow wealth and beneficence on the community, and singing kings' or warriors' praises ensured that the kings would live on in the memory of later generations. Fame lives on after death, and the concept had central importance in PIE society, especially for the warrior class. A phrase for 'imperishable fame' can be reconstructed for PIE on the basis of an exact equation between Sanskrit (*śrávas ákṣiti*) and Greek (*kléos áphthiton*, appearing in the excerpt from the *Iliad* in §12.65); in altered form the phrase appears also in several other branches. A warrior went into battle seeking fame because fame brought immortality, a way of overcoming death; a phrase for 'to overcome death' can be reconstructed for PIE, and survives ultimately in the Greek word *néktar*, the drink that bestowed immortality to the gods. (It is not inconceivable that the more militaristic aspects of the Indo-Europeans' successful spread owed something to this desire for achieving fame.)

2.38. The value placed on fame, and by extension on the poets that insured the immortality of a person's fame, is reflected in the generous largesse that poets received from their patrons in older IE societies. The relationship between poet and patron was mutually beneficial, one of reciprocal gift-giving: a king's or hero's livelihood in a very real sense depended on the preservation of his fame and on his reputation as surely as the poet's livelihood depended on being rewarded. A 'king without a poet' was proverbial in ancient Ireland for a poor king; and satire – the opposite of praise – was much feared and could have fatal consequences for the one at whom it was directed. The poet–patron relationship is neatly summed up in a Medieval Welsh account of a poet, Llywarch Hen, who, retreating from battle while carrying his slain patron's body, said, "I carry the head that carried mine." (In his artistic roles and his relationship to his patron, the IE poet is closely paralleled by the griots, poets and transmitters of traditional knowledge in western Africa.)

2.39. IE poetic tradition belongs to the type of poetry known as oral-formulaic poetry. Fundamental to this is the use of *formulaic* language, fixed words or groups of words that often had the function of filling out a verse-line. For example, in the *Iliad* Achilles is described as *pódas ókús* ('swift-footed'), a phrase that has a

convenient metrical shape; and the Homeric bards inserted it when the construction of a line needed a phrase of that shape, even in a passage where (for example) Achilles was sitting down. Having a storehouse of such formulae also makes it easier to compose and retell poems; although it is often stated that poems like the *Iliad* were memorized, in fact such poems are never the same each time they are retold because of on-the-spot improvisations and substitutions of one formula for another.

It is in the manipulation of formulae that IE poets showed their art, for IE poems are always a mixture of the old and the new; one hymn in the Rig Veda (3.31) has a passage that reads, "I make an anciently-born song new," while another (8.40) says, "Thus a new (poem) was spoken for Indra and Agni in the manner of the ancestors." Novelty was achieved first and foremost through the use of various grammatical, phonetic, and stylistic figures, some of which are described in the discussion to follow; in addition, an essential part of the aesthetic of much IE poetry was the use of obscure or difficult language. Skilled poets would often resist the temptation of using familiar formulae outright, but would distort them in certain ways, or allude to them obliquely with related notions or concepts. Examples of this are legion in the Rig Veda, the Old Avestan Gathas, the Homeric epics, and the odes of Pindar.

2.40. Formulae were more than just place-fillers; they had considerable cultural weight attached to them. Essential to the poet's ability both to use obscure language and to be understood was the manipulation of knowledge shared by his audience. This knowledge consisted in the associative semantic networks by which words and concepts were interconnected. Many of the major themes of IE poetry – religious truth, combat, heroic deeds, immortality, fame, reciprocity, fertility – found expression in certain words that had complex cultural and linguistic links to other words and concepts. Thus a poet could use one highly charged word or phrase as an allusion to a whole semantic complex, and by combining it with another word that brought to mind a different semantic complex could generate a new and often profound connection or equation, a cultural truth. The technique is not unlike the quotation of part of the melody of another work during a jazz improvisation: the brief quotation brings to mind the other piece and whatever associations that piece might have, and gives it new meaning in the context of the improvisation. Since part of the study of IE poetics necessitates figuring out these associative semantic networks, comparative IE poetics is crucially important for uncovering this most subtle component of PIE culture and world-view.

In the hands of a skilled poet, the density of culturally loaded verbiage can be staggering. To illustrate, consider the following excerpt from one verse in the Rig Veda (1.152.2):

satyó mántrah kaviśastá fghāvān

"True (is) the powerful formula pronounced by the poet." (trans. C. Watkins)

This snippet is only four words long, but each is packed with associations to important Vedic themes and formulas, many of IE antiquity, and their particular combination draws all those themes together in interesting ways. Let us begin with the final word, *fghāvān* 'powerful', literally 'possessing reproductive power'. It is very similar to another word, *ṛtāvān* 'possessing religious truth', which normally occupies the same position in a line; in the world of IE poetry, by its metrical position and

phonetic similarity *ṛghāvān* calls *ṛtāvān* and the latter's associations to mind. Religious truth (*ṛtā-*) is a central concept in Vedic India, and a primary vehicle of its expression was the poet's 'true formula' (*satyó māntrah*, the first two words of our line), which was conceived as bringing order to the universe. This formula is further described in our excerpt as *kaviśastā* 'pronounced (*śastā-*) by the poet (*kavi-*)'. The word *śastā-* 'pronounced, solemnly stated' is derived from a verb which has important connections of its own to the art of the poet. In particular, one of its derivatives, the noun *śāmsa-* 'praise', occurs in the famous compound word *nārāśāmsa-* 'praise of men', establishing an associative link between *kaviśastā-* and *nārāśāmsa-* and a sort of equivalency of *kavi-* 'poet' and *nār-* 'man, hero'. We know from elsewhere that this connection is real, and not a modern-day scholarly construct; compare the phrase *māntram nāryā ātakṣan* '(poets who) crafted the formula in a manly (*nāryā*) way' (Rig Veda 7.7.6; n.b. *māntra-* 'formula' here also!). And this brings us back to *ṛghāvān*, whose literal meaning 'possessing reproductive power' quite directly identifies the poet as virile; it, too, is a formulaic word in Vedic poetry and (together with closely related forms) is a traditional epithet of the warrior-god Indra, the head of the Vedic pantheon. Using *ṛghāvān* here to describe the praise-poem instead of the god being praised echoes the reciprocal relationship between gods and humans; this relationship was primarily expressed on the part of humans with a hymn of praise.

Metrics

2.41. A comparison of the metrics of the older IE poetic traditions reveals several different practices that, according to the most recent research, probably point to two distinct PIE poetic forms. The first, which has been recognized for a good century, had verse-lines of a fixed number of syllables and a rhythm that was quantitative, that is, based on a regular alternation of heavy and light syllables. (Light syllables are those ending in a short vowel; all others are heavy.) Lines came in longer versions of ten to twelve syllables and shorter versions of seven or eight syllables, and were grouped into strophes (stanzas) of three or four lines each. The longer lines had an obligatory caesura (break) neighboring the fifth syllable; both types were fairly free at the beginning but ended in a rhythmically fixed cadence (typical of many poetic systems around the world). The last syllable could be either long or short.

The following two strophes from the Rig Veda will illustrate these principles of long- and short-line construction. The symbol — represents a heavy syllable and v a light; the caesura is marked \parallel , and the last syllable is marked \times because it was indifferent to quantity (could be filled with either a light or a heavy syllable). The cadence is boldfaced; note that it is unchanging. The macrons indicate long vowels; o is also a long vowel; ṛ is a long syllabic r that counts as a long vowel; and th and bh are single consonants.

īndrasya nū vīriyāṇi prā vocam	$\text{— v v —} \parallel \text{— v — v —} \times$
yāni cakāra prathamāni vajrī	$\text{— v — —} \parallel \text{v v — v —} \times$
āhann āhim ānu āpas tatarda	$\text{v — v —} \parallel \text{v v — v —} \times$
prā vakṣāṇā abhinat pārvatānām	$\text{v — v —} \parallel \text{v v — v —} \times$

(1.32.1)

"I will now relate the manly deeds of Indra which he first did, wielding a cudgel: He slew the serpent, drilled through to the waters, (and) split the belly of the mountains."

mó šu varuṇa mṛṇmáyaṃ	— — — — — x
gṛhāṃ rājann ahām gamam	— — — — — x
mṛlā sukṣatra mṛlāya	— — — — — x

(7.89.1)

“Let me not go to the grave right soon, o Varuna, o king. Have mercy, o gracious ruler, have pity.”

It should be kept in mind that these were composed orally, long before writing came to India.

2.42. A second, potentially more archaic, poetic form has been argued for by Calvert Watkins and is termed the **strophic style**. Poetry in this form consists of strophes of relatively short lines whose structure is determined by grammatical and phonetic parallelism; there is no fixed line length or syllable count. It has also been termed, a bit misleadingly, as “rhythmic prose” (it is neither rhythmic nor prose). The strophic style is especially characteristic of archaic liturgical and legal texts, and certain mythological narratives. Some examples from several traditions follow:

(Hittite, Soldier’s Dirge)

Nešaš wašpeš Nešaš wašpeš
tiya-mmu tiya
nu-mu annaš-maš katta arnut
tiya-mmu tiya
nu-mu uwaš-maš katta arnut
tiya-mmu tiya

Shrouds of Nesa, shrouds of Nesa,
bind me, bind.
Bring me down for burial with my mother,
bind me, bind.
Bring me down for burial with my forefather,
bind me, bind.

(Avestan, Yasna Haptanḥāiti 35.4)

rāmācā vāstrəmācā dazdiiāi
surunuuatascā asurunuatascā
xšaiiaṇtascā axšaiiaṇtascā

(. . .) peace and pasture to be provided (by)
those who hear and those who do not hear,
those who rule and those who do not rule.

(Umbrian, Iguvine Tables VIa)

nerf arsmo
ueiro pequo
castruo fri
salua seritu
futu fos pace pase tua

(. . .) magistrates (and) formulations,
men (and) cattle,
heads of grain (and) fruits,
keep safe.
Be favorable (and) propitious in your peace.

(Armenian, The Birth of Vahagn)

erknēr erkin
erknēr erkir
erknēr ew covn cirani

Heaven was in labor,
Earth was in labor,
the purple sea also was in labor.

(Old Irish, “Cauldron of Poesy” §16)

Fó topar tomseo
fó atrab n-insce
fó comar coimseo
con-utaiḡ firse.

Good is the source of measuring
good is the acquisition of speech
good is the confluence of power
which builds up strength.

2.43. These examples exhibit most of the strophic style’s characteristic features. Grammatical parallelism and repetition is very frequent (“bring me down for burial

with my mother' ~ 'bring me down for burial with my forefather'; 'those who hear and those who do not hear' ~ 'those who rule and those who do not rule'; 'Heaven was in labor' ~ 'Earth was in labor'; 'Good is the acquisition of speech' ~ 'good is the confluence of power'). Bipartite alliterative phrases are also common (*salua seritu, futu fos, erknēr erkir, covn cīrani, topar tomseo, comar coimseo*). Also characteristic of the strophic style are bipartite phrases of various kinds that express a totality. One widespread type, called a **merism**, is represented by the Umbrian phrase *ueiro pequo* 'men (and) cattle', standing for the totality of movable wealth. The Avestan phrase *surunuuatascā asurunuuatascā* 'those who hear and those who do not hear' represents a second type; compare, from the Latin prayer in §13.53, *morbos uisōs inuisōsque* 'diseases seen and unseen'. Yet other types are represented by Old Persian *hašiyam naiy duruxtam* 'true and not false' and Greek *litās t' epaoidās* 'prayers and incantations' (two synonyms or near-synonyms).

2.44. Many IE poetic forms seem to combine fundamental features of the strophic style (such as alliterating word-pairs and freedom in the number of syllables) with some structural rigidity reminiscent of the rhythmic/quantitative forms (a fixed rhythm at line-end preceded by a caesura, or a fixed number of syllables for the whole line). This is perhaps clearest in poetry like the following South Picene epitaph (see §13.73):

postin viam videtas	Along the road you see
teris tokam alies	the toga (?) of Titus Alius (?)
esmen vepses vepeten	buried (?) in this tomb.

It can be divided either into bipartite alliterating phrases (excluding the first word) or into three seven-syllable units ending in a trisyllable. Such seven-syllable lines with internal alliteration and a final trisyllable are found also in Luvian and Irish, and a similar eight-syllable line ending in a trisyllabic cadence (but without alliteration) is characteristic of traditional Slavic historical ballads. Lines ending in a trisyllabic cadence and with a variable number of syllables before are characteristic of archaic Irish verse and of archaic Roman Saturnian verse. In traditional Germanic heroic verse, a line is divided into two half-lines, the first of which has two words that alliterate and that further alliterate with one word in the second half-line; there is some freedom in the number of syllables, but the first half-line in particular tends to be weighted toward the end (but without a fixed cadence).

Repetition of sounds (including alliteration, assonance, and, less frequently, end-rhyme) is characteristic of IE poetry even outside the strophic style. A line like the following, from the Roman comic playwright Plautus (*Miles Gloriosus* 603), is quite typical of the technique:

sī minus cum cūrā aut cautēlā locus loquendī lēctus est
 "If your place of conference is chosen with insufficient care or caution . . ."
 (trans. P. Nixon)

We have the alliterating *k* sounds (spelled *c*) of *cum cura aut cautela* followed by *l*'s in *locus loquendi lectus*, all of which also have *k* sounds in their interior (*locus loquendi lectus*). The two words *cum cura* both have *u*'s, and the following two words *aut cautela* share the sequence *aut*; and note the repeated *oc oc ec* of *locus loquendi lectus*. In Plautus, the repetition of these sounds is partly for comic effect;

in a line like the following, from a different tradition (*Iliad* 11.547), the phonetic figures underscore the sense and add gravity:

entropalizómenos olígon gónu gounòs ameibôn
 "continually turning his head (and) shifting one knee past the other a little"

At this moment in the story, the Greek warrior Ajax is being struck with fear, and the repeated *gon gon goun* is iconic of his jittery demeanor. (The earlier form of *gounos* was *gonwos*, which was probably the form used when the line was composed; so originally the line contained a perfectly repeating *gon gon gon*!)

BEHAGHEL'S LAW

2.45. Among the many other poetic techniques that IE poets availed themselves of, mention may be made of a tendency to give more verbal flair to the last thing enumerated in a series, a practice called Behaghel's Law of Increasing Members, after the Germanicist Otto Behaghel. Typical examples include, from Sanskrit, *Damaṃ Dāntaṃ Damanaṃ ca suvarcasam* 'Dama, Dānta, and Damana having good life' (*Mahābhārata* III 50.9); from Old English, *Heorogār ond Hrōðgār ond Hālga til* 'Heorogar, Hrothgar, and Halga (the) good' (*Beowulf* 61); and a double example from Greek, from the Catalogue of Ships in the *Iliad* (2.532–3), *Bēssān te Skárphēn te kai Augeiās erateinās Tárphēn te Thróniōn te Boagriou amphi rhēethra* '(who lived in) Bessa and Skarphe and beautiful Augeiai, and Tarpe and Thronion and along the waters of Boagros'.

Personal Names

2.46. In a society where the spoken word was of such importance, it is no surprise that bestowing a name upon a newborn was the subject of a ritual. We can reconstruct the phrase for the act itself, literally 'make a name', on the basis of such cognate phrases as Hittite *lāman dāi-*, Vedic Sanskrit *nāma dhā-*, Greek *ónoma tithesthai*, Lat. *nōmen in-dere*, and Tocharian A *ñom tā-*. In Vedic India, the name is given as part of the *nāmadheya-* ('name-placing') festival on the tenth day after the mother has left the childbed and been bathed; in Greece the name was given on the tenth day also. In Rome, the name was given nine days after the birth of a boy, eight days after the birth of a girl. Among Germanic peoples, according to Alemannic and Frankish legal texts the name was given on the ninth or tenth day after birth, accompanied by a ritual; here too the mother was first bathed. These facts would indicate that in PIE society, the mother recovered for nine (?) days after childbirth, rose, was bathed, and the child was then named.

2.47. The importance attached to names is due partly to the connection between one's name and one's reputation, especially among the ruling or warrior classes. In several IE traditions, the inherited words for 'name' and 'fame' are collocated (e.g., Greek *onomáklutos* 'famous in name', Tocharian A *ñom-klyu* 'name-fame', Vedic Sanskrit *śrútyam nāma* 'famous in name'). In Old Irish, an everlasting name was synonymous with everlasting fame.

2.48. We have a large dossier of ancient Indo-European names, many of which furnish information about naming practices in PIE times, especially (again) within

the warrior class. One very common type of name, found in most branches and securely reconstructible for PIE, is a bipartite compound X-Y where one or both compound members are concepts, virtues, or animals that were important in Indo-European society, such as fame, guest, god, and strength. A sampling will illustrate their character:

'fame'	Old Russian <i>Bole-slavŭ</i> 'having greater fame/glory' Illyrian <i>Ves-cleves</i> 'having good fame' Greek <i>Themisto-klēs</i> 'law-fame'
'guest'	Lepontic <i>Uvamo-kozis</i> 'having supreme guests' Runic <i>Hlewa-gastiz</i> 'fame-guest'
'protection'	Luvian <i>Tarhunta-zalmaš</i> 'having the Stormgod as protection' Old High German <i>Ans-elm</i> 'having god as helmet' Gaulish <i>Anextlo-mārus</i> 'great in protection'
'god'	Vedic <i>Devā-śravās</i> 'having divine fame' Czech <i>Bohu-slav</i> 'having the fame of god' Greek <i>Dio-génēs</i> 'born of god, born of Zeus' German <i>Gott-fried</i> 'having the peace of god'
'battle'	Gaulish <i>Catu-rix</i> 'battle king' Old Welsh <i>Cat-môr</i> 'great in battle' Old High German <i>Hlūd-wīg</i> 'loud in battle' (<i>Ludwig</i>)
'people'	Greek <i>Agé-lāos</i> 'leader of the people' Old High German <i>Liut-pold</i> 'brave among the people' (<i>Leopold</i>) Gothic (Latinized) <i>Theode-ricus</i> 'people's king'
'man, hero'	Irish <i>Fer-gus</i> 'having a hero's strength' Old Persian <i>Xšay-āršā</i> 'hero among kings' (<i>Xerxes</i>) Greek <i>Aléks-andros</i> 'warding off heroes' (<i>Alexander</i>)
animals	Vedic <i>Ṛjī-śvā</i> 'having swift dogs' Old Norse <i>Ráð-ulfr</i> 'counsel wolf' (<i>Ralph</i>) Gaulish <i>Mori-tasgus</i> 'sea badger' Ogam Irish <i>Cuno-rix</i> 'dog king'

2.49. In several IE societies, names of sons were picked (or created) so as to resemble the names of their fathers in specific ways. This was often done by recycling one of the compound members: Greek *Dīno-krātēs* 'having fearful strength', son of *Dīno-klēs* 'having fearful fame'; *Eu-krātēs* 'having good strength', son of *Euru-krātēs* 'having broad strength'; Old High German *Walt-bert* 'bright in power', son of *Wald-ram* 'power raven'; *Hilti-brant* son of *Haðu-brant*, both meaning 'battle sword'. Such practices are partly behind the appearance of nonsensical compound names like Old High German *Fridu-gundis* 'peace-battle' or Greek *Rhód-ippos* 'rose-horse', where one of the compound members was copied over from the name of one's father without regard for what the new compound would mean.

Archaeology and the PIE Homeland Question

2.50. No issue in Indo-European studies has aroused more controversy or popular interest than determining the place the Proto-Indo-Europeans called home. The variety of different suggestions put forth – some of them quite outlandish – is

enormous. As J. P. Mallory writes in his book *In Search of the Indo-Europeans* (p. 143),

We begin our search for the homeland of the Indo-Europeans with the deceptively optimistic claim that it has already been located. For who would look further north than [C]ok[a]manya Tilak and Georg Biedenkapp who traced the earliest Aryans to the North Pole? Or who would venture a homeland further south than North Africa, further west than the Atlantic or further east than the shores of the Pacific, all of which have been seriously proposed as cradles of the Indo-Europeans? This quest for the origins of the Indo-Europeans has all the fascination of an electric light in the open air on a summer night: it tends to attract every species of scholar or would-be savant who can take pen to hand. It also shows a remarkable ability to mesmerize even scholars of outstanding ability to wander far beyond the realms of reasonable speculation to provide yet another example of academic lunacy.

Allied with the purely geographical question of where the Proto-Indo-Europeans lived is the temporal question of when they lived there. The insurmountable difficulty with answering both these questions is the simple fact that no material artifact of a preliterate people, nor their mortal remains, can tell us what language they spoke. Much of the time, we cannot even assuredly identify a type of artifact with a particular people: styles of pottery, for example, can diffuse from culture to culture, just as computer technology has spread globally today. While we can reconstruct a number of helpful terms relating to PIE material culture (see the next section), none of them is specific enough for matching particular archaeological finds with speakers of PIE. For example, we can reconstruct a word for 'fortification', but we are in the dark about whether there was an "IE type" or style of fortification that could be identified with particular prehistoric fortifications of Eurasia. As Mallory rather sardonically puts it (p. 126), "Indeed, it is bizarre recompense to the scholar struggling to determine whether the Proto-Indo-Europeans were acquainted with some extremely diagnostic item of material culture only to find that they were far more obliging in passing on to us no less than two words for 'breaking wind'."

Nevertheless, the situation is not entirely hopeless, and a careful consideration of the linguistic and the archaeological record does provide a number of important and tantalizing clues. Let us first consider the question of when PIE was spoken, for which we must review what we know about the Indo-Europeans' material culture and technology.

Material culture and technology

WHEELS AND TOOLS

2.51. The Proto-Indo-Europeans knew the wheel, for which they had at least two words, one of which (the family of Eng. *wheel*, Gk. *kúklos*, and Sanskrit *cakráṃ*) is found in most of the branches (see further §2.58 below). We can also reconstruct words for wheel hub (nave) and axle, and a specific verb referring to the act of conveyance in a vehicle; from this verb the noun for 'wheeled vehicle, wagon' was derived. The use of draft animals for pulling the wheeled vehicles required yokes and thills (yoke poles); terms for both in PIE have been reconstructed.

PASTORALISM

2.52. The Proto-Indo-Europeans practiced agriculture and made use of various farming implements. A verb meaning 'to plow' is securely reconstructed, and several branches have similar words for 'plow' (the implement) that are probably inherited. We also know the words for some other farming tools, such as the harrow and sickle. Although words for grain, for threshing and grinding grain, and for some specific grains can be reconstructed (wheat, barley, and probably emmer and spelt), it is uncertain whether grains were cultivated by the Proto-Indo-Europeans; however, such cultivation is strongly suggested by the fact that grains have a prominent role in the mythology, folklore, and ritual practices of many IE traditions that can be projected back onto the proto-culture and that point to the importance of cereals for their livelihood. The PIE word for 'field' has descendants in most branches, and was a derivative of an equally widely represented verbal root referring to leading or driving cattle, which points to the use of draft-oxen in plowing. Slavic, Germanic, and Celtic have cognate words for 'fallow', indicating that their ancestors may have engaged in shifting cultivation; but we do not know if this is an inheritance from PIE, as it may also be a later, locally innovated term of these three geographically contiguous branches.

2.53. As we have seen, an important part of the Indo-Europeans' material culture consisted in stockbreeding, and we are well informed about the IE words for various domesticated animals. Among bovines, the PIE lexicon distinguished cow, steer, ox, and bull. The Indo-Europeans were familiar with dairy products (whether from cows, goats, or mares): we know their words for coagulated or sour milk, butter, and curds. Interestingly, while we can reconstruct a verb for the act of milking, we cannot reconstruct a word for the liquid itself: the terms in the daughter languages are apparently related (e.g. Gk. *gálakt-*, *glak-*, Lat. *lact-*, Eng. *milk*), but differ too much from one another to allow precise reconstruction.

2.54. To these may be added words for sheep (and its wool), ram, lamb, goat, horse, and dog, the last of these domesticated earlier than any of the others. We do not know if ducks and geese were domesticated, but we can reconstruct the words for them. There were separate roots for piglet and fully-grown pig, suggesting that swine were domesticated (the two are treated differently in animal husbandry). The horse is often thought of as the IE animal par excellence; it was important in PIE myth and ritual (see §§2.23 and 2.26 above), and is thought to have played a critical role in the rapid expansion of the early IE tribes, on which more presently.

DWELLINGS

2.55. Some terms pertaining to houses and house-construction have survived. The general PIE word for house has descendants in most branches, as does the word for door (usually attested in the dual or plural; presumably doors came in pairs). We can also reconstruct a word for doorjamb. Doors were probably kept shut with pegs of some kind, referred to by a word whose descendants variously mean 'key', 'peg', or 'nail' and that is derived from the verb meaning 'to close'. Roofs were thatched; words for 'roof' in some IE languages are cognate with words for 'thatch' in others, all of them derived from a verbal root meaning 'to cover'. Inside a dwelling was the hearth, which had great symbolic and even religious significance (recall §2.25 above) and for which a PIE word can be reconstructed.

FOOD PREPARATION

2.56. At least four branches – Indo-Iranian, Greek, Italic, and Germanic – have cognate words for oven. Cooking, baking, and boiling are all terms we can reconstruct, as is a term for broth. Grinding grain could be done in a hand-mill or *quern* (the direct descendant of the PIE term for this implement). Also part of the PIE culinary dossier was fermentation: the Indo-Europeans drank mead, their word for which has descendants in most of the daughter branches. Whether wine was known is a contentious issue. Viticulture is at least as old as the sixth millennium BC in the Caucasus, and the word for ‘wine’ is the same in the IE family (Eng. *wine*, Lat. *uīnum*, Gk. (*w*)*oīnos*, Russ. *vin*o, etc.), Semitic (**wayn*-), and the Kartvelian languages of the Caucasus (e.g. Georgian *ghvino*). Though some believe the word is native PIE, the arguments for this are speculative, and most researchers believe rather that it diffused into the IE languages at a post-PIE date.

TEXTILES AND CLOTHING

2.57. The well-represented word for sheep’s wool stood alongside a word for linen (or flax); these, together with reconstructed roots for sewing, spinning, weaving, and plaiting, and nouns for needle and thread, show that the Proto-Indo-Europeans produced textiles. A verbal root meaning to clothe has descendants in most of the branches; we can also reconstruct the verb for girding, which formed a derivative noun for belt. The Proto-Indo-Europeans were also familiar with combs, and with ointments or salves. Aside from this, though, we know little about their dress or bodily adornments.

2.58. This represents the bulk of the linguistic evidence for the Indo-Europeans’ material culture. There are various additional terms, of course, as for other tools and weapons, but the latter are archaeologically widespread and go back much further than most dates entertained for PIE. A prehistoric society somewhere in Eurasia that practiced agriculture and stockbreeding, had the plow and other specialized agricultural implements, and had developed secondary products from milk and wool would most likely have belonged to the late Neolithic (fifth and fourth millennia BC). Based on the available archaeological evidence, the addition of wheeled vehicles to this picture allows us to narrow the range to the mid- or late fourth millennium: the earliest wheeled vehicles yet found are from c. 3300–3200 BC. If one adds a century or two to that figure (on the assumption that the actual invention of wheeled vehicles predates the earliest extant remains), that means the latest stage of common PIE (the stage directly reachable by reconstruction and before any of the future branches separated) cannot have been earlier than around 3400 BC. The wheel- and wagon-related terms are particularly telling linguistically: all of them save ‘thill’ are transparently secondary, that is, derived from known roots. One of the words for wheel, the ancestor of Sanskrit *cakrām*, Gk. *kúklos*, and Eng. *wheel*, is derived from the verb ‘to turn’ and has the look of a colloquial neologism, as we will discuss in §6.63. The other word, represented e.g. by Latin *rota*, is from the verb ‘to run’. ‘Nave’ is identical to, or related to, the word for ‘navel’, and ‘axle’ is also the word for ‘shoulder joint’. ‘Yoke’ is derived from the verb ‘to join’. Thus the whole complex of terms looks like a set of new metaphorical extensions of already existing terms to denote novel technologies. (Compare the extension of *mouse* to denote a computer input device.)

2.59. We may add to this the IE lexical facts concerning metals. Iron and tin do not appear in the archaeological record until after the fourth millennium, while copper is found already by the early sixth millennium, and silver in the late fourth millennium. Aside from a general word for 'metal' (which may, for all we know, have referred to copper), the only specific metal term that has any chance of being of common PIE date is 'silver' (represented e.g. in Avestan *ərəzātəm*, Latin *argentum*, and Gaulish *arganto-*). Although it is a transparent derivative of an adjectival root meaning 'shiny, white', in five of the six branches where it appears it has the same unusual morphology, making common descent from PIE more likely than not. This is consistent with a breakup of PIE in the late fourth millennium.

Linguists would not be comfortable with a date much later than this; by common consent a later date would have left insufficient time for Anatolian and Indo-Iranian to diverge as much as they had diverged by the time their first written traces appear in the early and mid-second millennium. Note, though, that there is no generally accepted method for determining average rates of language change, and it is far from assured that such rates even exist. But there are other reasons to think that PIE had broken up by the end of the fourth millennium, which we will look at shortly.

Location of the homeland

2.60. Several methods of approach have been used for locating the homeland of the Proto-Indo-Europeans. Most famously, perhaps, are investigations of IE words for the natural world – the physical environment, fauna, and flora, in the hope that one of these terms would refer to something with a very limited geographical distribution. The reconstructed terms for topographic features are too basic to be helpful: mountain, river, lake or sea, and marshy land. These do no more than rule out a desert home for the Indo-Europeans (as does the reconstructed word for 'boat!'). The terms for animals and plants, however, are more varied and have aroused greater interest.

FAUNA

2.61. We can reconstruct the names for the bear, fox, wolf, beaver, otter, hedgehog, and elk (or deer), as well as words for quite a few birds, including the sparrow, quail, thrush, crane, vulture, blackbird, crow, raven, eagle, falcon (or hawk), jay, kite, pheasant, stork, and probably owl. (Some specialists dispute one or another of these.) As for reptiles and aquatic animals, we only know terms for the turtle, frog, and snake, plus two general words for fish and terms for a few specific fish species, especially the trout and salmon or a related species (see further below), and a word for some large fish that later came to be applied to cetaceans. Celtic and Indo-Iranian have words for leech that are probably cognate, and two rhyming words for worm are widely represented. As for insects, we know the name of the bee, together with words for its honey and wax, as well as the PIE words for wasp and hornet. The Indo-Europeans were also well acquainted with those perennial unwanted companions of humans and their dwellings, the mouse and louse (a rhyming pair then as now); reconstructible alongside the latter is the word for its egg, the nit, and a word for flea.

2.62. Almost all these animals are ubiquitous throughout Europe and large parts of Asia, with the possible exception of the beaver, which is not found in Greece or Anatolia, locations that are doubtful for the PIE homeland for many other reasons

(though see §2.71 below). But these terms do suggest that the Indo-Europeans lived in regions that were at least partly forested. The bear and elk are forest-dwelling animals, as are honeybees, and the otter and beaver dwell specifically in riverine forests. Some excitement was generated at one point by a theory that one of the reconstructed fish-names, **loksos* (the ancestor of English *lox*), referred specifically to the Atlantic salmon (*Salmo salar*), whose distribution in Eurasia is limited to rivers flowing into the Baltic. But it is now believed that the word referred to the salmon trout (*Salmo trutta*), which is found over a much wider area. (Whatever the **loksos* was, it seems to have been viewed as pre-eminent, to judge by the fact that it became the general word for 'fish' for the Tocharians. On the other hand, this may not mean anything, given that the PIE word for 'louse' became the general Tocharian word for 'animal'!)

FLORA

2.63. The reconstruction of names for plants is more problematic than for animals. Very few plant names in the western IE languages have cognates in the eastern languages, a fact that has occasioned much debate over how many terms to reconstruct for PIE. The most clear-cut is the birch, whose name is found in six branches; after it come the willow and ash (the latter the preferential wood for making spears), and it is possible though more controversial to include the pine and yew. The oak was known, since the word for 'acorn' is securely reconstructible for PIE (though the most widespread word for 'oak' is only found in the European branches). These trees occur pretty much all over Europe and Asia. Hittite and Irish preserve a cognate word for 'hawthorn', and also attest to magical uses of it that may be inherited from PIE. Five branches attest a word that sometimes means 'beech' (e.g. *beech* in English), sometimes 'elder' (e.g. Russian *buzina*), and 'oak' in Greek *phēgós*. For a while, attempts were made to show that the word meant 'beech' originally, because it used to be thought that the beech was found only west of Russia. However, we now know the range of the beech extended much farther east in the recent past, so the whole issue is moot for determining the location of the PIE homeland.

2.64. The European languages allow the reconstruction of words for such trees as the oak, elm, juniper, alder, apple, hazel, and cherry. These may have been PIE, or specifically European terms that arose later; complicating our analysis is the fact that several of these words, such as 'apple' (**abel-*), have an un-Indo-European look to them that suggests borrowing (although the borrowing could itself be of PIE date). See further §4.11 on this. Greek, Latin, Germanic, and Slavic share a word variously referring to mistletoe, black cherry, or birdlime; birdlime can be obtained from both mistletoe and black cherries, but which plant was designated by the term in PIE is uncertain. We also know the PIE words for berry and bean. All of these flora are quite widespread, but some of them are not found very far south, suggesting (together with the presence of a PIE word for 'snow') a temperate rather than a tropical or subtropical region.

BURIALS AND THE KURGAN CULTURES

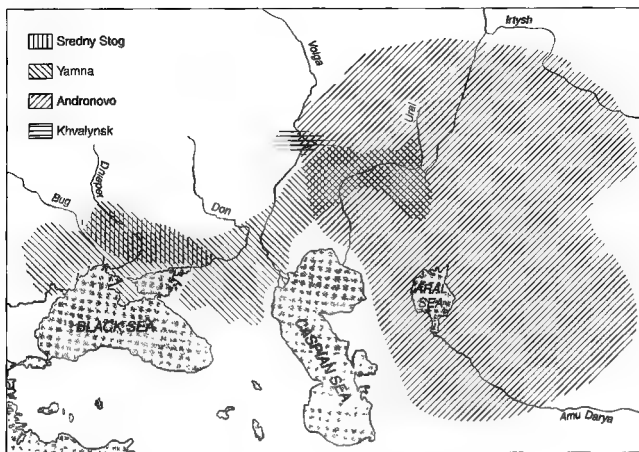
2.65. Bridging the gap between material and non-material culture, and of particular importance in archaeology, are burials. While a culture might well adopt technological innovations from other cultures, burial practices are considerably more

resistant to outside influence. Ancient IE texts describing burials, especially of kings or warriors, are known from several branches, and although they do not agree with one another in every detail, they allow us to piece together a reasonably good picture of PIE burial practices. A dead person was buried in his own individual tomb that was like a mortuary house and heaped over with earth (a *tumulus* or burial mound). The corpse was sometimes cremated; this was the norm in the Indo-Iranian world and a special honor for heroes in ancient Scandinavia. Buried with the deceased were various grave goods, including ornaments, food, clothing, weapons, tools, and often wheeled vehicles, sacrificed animals, and even people. All these things would be needed in the afterlife.

2.66. Various Copper and Bronze Age cultures in the steppeland of southern Russia, around the Black Sea and middle Volga, are associated with characteristic tumuli called *kurgans* (from the Russian term). The *kurgans* and the burials they contain are consistent with the early IE burial practices outlined above, and the late Lithuanian archaeologist Marija Gimbutas proposed that the *kurgan* peoples were in fact early Indo-Europeans. The archaeological excavations, in Gimbutas's view, indicate that the *kurgan* cultures had a pastoral economy, hierarchical social structure, patriarchy, aggressive warfare, animal sacrifice, worship and/or use of the horse, wheeled vehicles, and worship of a solar deity. All these are Indo-European cultural characteristics.

Gimbutas emphasized the differences between *kurgan* burials and the indigenous "Old European" burials that predate the intrusion of the Indo-Europeans into Europe. The "Old European" burials were in oval or egg-shaped tombs, without the sorts of grave goods seen in *kurgan* burials. Gimbutas viewed this "Old European" culture as *matrifocal* (that is, having a ritual focus on women and goddesses), peaceful, and goddess-worshipping – strongly contrasting with the patriarchal, aggressive, skygod-worshipping Indo-Europeans. While her views have engendered considerable controversy, it is fair to say that one or another version of the *kurgan* theory is accepted by most archaeologists and Indo-Europeanists today, and the differences between the *kurgan* cultures and the "Old European" cultures are quite plain.

2.67. The term "*kurgan* culture" has been criticized as too vague; but if we start to look at some more specific cultural entities within the general area of the Pontic-Caspian steppes (the vast grasslands north of the Black and Caspian Seas), a coherent historical picture emerges of considerable interest. Recall from the previous discussion that a date of c. 3300–3400 BC – the invention of wheeled vehicles – was the earliest possible date for the breakup of common PIE. A date of about 3500 BC corresponds with the first large-scale occupation of the Pontic-Caspian steppes, by a people known as the *Yamna* culture (or *Yamnaya* culture). This culture originated in the borderland between the Pontic-Caspian steppes and the neighboring forest regions between the Dnieper and Volga Rivers. (These dates – 3500 and 3300–3400 BC – can be taken as archaeologically equivalent, for carbon-14 dating has a range of error of two to three centuries.) From as early as 3500 BC comes the first possible (though very controversial) evidence of horseback riding in the archaeological record, in the form of microscopic abrasions on horses' teeth from clamping down on a bit. This evidence of bit microwear, as it is called, is primarily associated with the related *Botai* culture, as well as probably with the *Yamna*. The discovery was made by the archaeologists David Anthony and Dorcas Brown, and if their interpretation is correct, it pushes horseback riding much further back than previously thought. (The oldest pictorial representations of horseback riding date only to about 2000 BC.)



Map 2.1 Selected Late Neolithic and Chalcolithic (Bronze Age) cultures north of the Black and Caspian Seas

2.68. This has enormous consequences for the whole question of the IE homeland and expansions. Anthony and his colleagues have emphasized that the advantages lent by horseback riding are far more than just military, especially for a people who had previously been confined to riverine forested regions for their livelihood. Horseback riding would have allowed the population to scout far and wide for new pastures, transport goods quickly, undertake large-scale livestock breeding and herding, sustain a mobile and flexible pastoral economy, and engage efficiently in long-distance trading (as well as raiding and warfare). There is archaeological evidence for all of these activities on the part of the Yamna, and they were the first people in the Pontic-Caspian area to spread into the deep steppe and exploit it.

2.69. Importantly, the Yamna can be linked rather clearly with a later cultural complex that we are reasonably sure was Indo-European – specifically, Indo-Iranian: the Andronovo culture, the earliest archaeological complex that can be identified with a particular IE linguistic group. The Andronovo culture evinces numerous features specifically described in early Indo-Iranian texts, especially aspects of tomb-construction and burial ritual. Their kurgan burials, some of them spectacular, contain wheeled vehicles, livestock (horses, sheep, goats, cattle), weapons, ornaments (including cheek-pieces for horses), and scatterings of sacred straw (called *barhis* in the Rig Veda). The Andronovo appeared around 2200 BC in northern Kazakhstan between the Ural and Irtysh Rivers, and is generally agreed to have ultimately developed from an early third-millennium eastern outgrowth of the Yamna.

2.70. We can therefore follow a direct line back from early Indo-Iranians to a cultural complex of the Pontic-Caspian steppes of the mid-fourth millennium BC.

Does this mean the Yamna culture was in fact the early Indo-Europeans? Some see the Yamna as merely ancestral to the Indo-Iranians. But given that we have already established a date of c. 3400 BC as the earliest possible end of PIE linguistic unity, it seems quite unlikely that anything specific enough to be identified as Indo-Iranian or pre-Indo-Iranian was already on the scene at that time. The Yamna culture, in fact, certainly fits the bill of being the late Proto-Indo-Europeans: rapid expansion from an original area that comprised a temperate forest; the recent adoption (or invention) of wheel technologies; widespread stockbreeding; and use of the domesticated horse, including in ritual.

2.71. However convincing this scenario just outlined may be, its tentative and controversial nature cannot be too strongly emphasized. Part of it hinges crucially on the date of the domestication of the horse, and of horseback riding in particular. Both of these are hotly debated issues, although more and more scholars agree that the horse was at least domesticated by the time of the Yamna. We do not have any written texts to prove or disprove the late fifth- and early fourth-millennium-BC Pontic-Caspian steppes as the IE homeland, and will likely never have any. There is thus ample room for other views.

Probably the most prominent alternative hypothesis is that the homeland was in Anatolia (eastern Turkey) and neighboring regions in the Caucasus, as most recently propounded by Colin Renfrew of Cambridge University and, in a somewhat different version, by Tamaz V. Gamkrelidze and Vyacheslav V. Ivanov. This view has not been widely accepted within IE studies. Opponents point to the lack of archaeological evidence for an Anatolian homeland or for early dispersal from Anatolia; by all indications, Indo-European-speaking peoples only came into Anatolia rather late. Renfrew places the IE homeland in Anatolia as part of a much broader theory concerning the nature of the expansion of the Indo-Europeans, which he sees as gradual and peaceful rather than aggressive. This theory cannot be treated here in detail, but it is beset with a whole host of problems, not the least of which is that it requires a fantastically early date for the breakup of PIE – a full three millennia before the earliest known wheeled vehicles. This date can only be maintained by willfully ignoring the comparative linguistic evidence discussed earlier in this chapter.

2.72. By contrast, several considerations, both archaeological and linguistic, point independently to the Yamna of the Pontic-Caspian steppe as the Indo-Europeans in their early expansionist phase. The Yamna itself is usually agreed to have developed partly out of the *Sredny Stog* culture, which stretched over an area north of the Black Sea with a western boundary around the Dnieper River, and out of the *Khvalynsk* culture, which was located east of the *Sredny Stog*. The area inhabited by the *Sredny Stog* has been seen by some as the “real” PIE homeland. The *Sredny Stog* flourished c. 4500–3500 BC, until the appearance of the Yamna. There is no conclusive evidence for horseback riding from this period; horses were apparently mainly used for hunting and for food, and settlements were permanent rather than nomadic.

EUROPE AND THE INDO-EUROPEANS

2.73. It is more difficult to connect the various European cultures with the Pontic-Caspian; in fact, the archaeological prehistory of all the European IE groups is quite murky. But in the period 3100–2900 BC came a clear and dramatic infusion

of Yamna cultural practice, including burials, into eastern Hungary and along the lower Danube. With this we seem able to witness the beginnings of the Indo-Europeanization of Europe. By this point, the members of the Yamna culture had spread out over a very large area and their speech had surely become dialectically strongly differentiated; a common PIE must no longer have existed. Though their initial spread by around 3500 BC may have predated the introduction of wheeled vehicles (or it may not, given the range of error in carbon-14 dating mentioned above), they had apparently not dispersed far and wide enough, or become dialectally differentiated enough, to prevent the spread of a unified wheel-related vocabulary from one part of the speech community to the rest of it.

So the search for the Indo-Europeans goes on. Whether we are closer to finding them now, or whether we are just as far from the truth as ever, is perhaps not knowable. As with the reconstruction of the Indo-European proto-language that will occupy our attention in the next several chapters, and as is true in the sciences in general, we can only deal in probabilities – what explanations have the best and most convincing evidence and arguments to back them up. Yet this is no mere academic exercise. The Indo-Europeans have been a uniquely successful people, whose impact on human history has been as great as, if not greater than, that of any other. We will not be able to understand the historical fact of their spread and success without first illuminating their origins.

For Further Reading

The classic ethnolinguistic investigation of PIE vocabulary is Benveniste 1969, a collection of over fifty studies covering livestock and wealth, economic obligations, kinship, social status, law, and religion. It is a unique work, summing up a lifetime of research into IE culture by one of the twentieth century's foremost Indo-Europeanists. The standard dictionary of PIE is Pokorny 1959–69, in German; the only English work of similar scope is Mann 1984–7, which however is substandard. Smaller but very useful, especially for English-speaking non-specialists, is Watkins 2000, which contains all the Indo-European roots (over 1300) that have descendants in English, with entries detailing the intermediate prehistoric stages. The volume begins with an excellent overview of PIE culture and vocabulary. Also of high usefulness is Buck 1949, which lists groups of terms in most IE branches by category, with etymological commentary on each group.

A recent reference work on IE vocabulary, culture, and archaeology is Mallory and Adams 1997, which contains hundreds of articles, some fairly extensive, by numerous contributors. The quality and coverage of the articles are uneven, but it is still a welcome volume and particularly useful for the archaeological information it provides. Also recommended are the technical articles contained in the second volume of Watkins 1994 and the second volume of Gamkrelidze and Ivanov 1984; the latter contains an extensive collection and discussion of cultural and folkloric material.

Good introductions to the theories of Georges Dumézil are Dumézil 1958 and the slightly earlier Dumézil 1952. Contemporary studies in the Dumézilian mold frequently grace the pages of the semiannual *Journal of Indo-European Studies*, which also publishes many articles on IE archaeology and mythology. Puhvel 1987 is a lively and useful comparative collection of various IE myths.

For IE poetics, global reference may be made to Watkins 1995, which is an exhaustive study of the IE dragon-slaying myth and of IE poetic practice. Readers of German and Italian should also consult Schmitt 1967 and the writings of Enrico Campanile, such as Campanile 1977.

The best recent overview of IE archaeology and the attempts to locate the PIE homeland and reconstruct PIE culture is Mallory 1989, with copious maps and illustrations; it is very accessible to a general audience. Mallory's command of the linguistic issues is better than most archaeologists'. A useful collection of articles by Marija Gimbutas on these and related subjects is her collection Gimbutas 1997. Renfrew's alternate theory discussed in §2.71 is presented in Renfrew 1987.

For Review

Georges Dumézil	hospitality	'imperishable fame'	strophic style
trifunctionalism	Father Sky	oral-formulaic poetry	merism
Omaha kinship	divine twins	formula	Behaghel's Law
system	<i>aśvamedha</i>	quantitative verse	kurgan
<i>Männerbund</i>	taboo deformation		Marija Gimbutas

Exercises

- 1 Write a sentence or two describing or exemplifying each of the following aspects of PIE culture:

a social stratification	f enforcement of justice	k outlaws
b patrilineal exogamy	g gods and humans	l poets
c movable wealth	h consecration of kingship	m fame
d gift and counter-gift	i medical treatment	n burial customs
e guest-friendship	j dragon-slaying myth	

- 2 Briefly explain the significance of each of the following items for attempts to locate the homeland of the Proto-Indo-Europeans:

a salmon	b burial customs	c beaver	d beech
----------	------------------	----------	---------

- 3 Briefly explain the significance of each of the following items for establishing the approximate date that PIE was spoken:

a wheeled vehicles	b milk products	c silver	d plow
--------------------	-----------------	----------	--------

- 4 Briefly describe or explain the following and, where relevant, comment on their significance:

- a the PIE phrase **(e-)gʷhen-t ogʷhim*
- b the journey of the soul to the afterlife
- c the relationship between poet and patron
- d to 'lead away' a woman in marriage
- e assonance
- f name and fame
- g horseback riding

- 5 Below is an excerpt from Homer's *Odyssey* (19.106–118). Odysseus has made it home to Ithaca but has not revealed himself to his wife Penelope yet. Here is his first encounter with her, while he is still in disguise as a beggar. The Greek (slightly simplified) is given first, followed by an English translation; the various boldfacings, italicizings, etc. are designed to help you match the Greek words with the words in the translation.

Tên d' **apameibomenos** *prosephê* polumêtis Odu~~ss~~eus:

"ô gunai, **ouk an tis se** **bro~~t~~ôn** ep' apeirona gaian

neikeoi; ê gar **seu** kleos ouranon eurun *hikânei*,

hôs te teu ê basilêos amûmonos, *hos te theoudês*

andrasin en polloisi kai iphthîmoisin *anassôn*

eudiklâs anekhêisi, pherêsi de gaia melaina

pûrou kai krîthas, *brithêisi de dendrea karpôi*,

tiktêi d' empeda **mêla**, thalassa de **parekhai** ikhthûs

ex **euêgesiês**, *aratôsi de laoi hup' autou*.

tôi eme nûn ta men **alla** **metalla** *sôi eni oikôi*,

mêd' emon exereine genos kai **patrida** *gaian*,

mê moi mallon **thûmon** eniplêsêis *odunâôn*

mnêsamenôi: . . .

110

115

Answering her, wily Odysseus said:

"O lady, no mortal on the boundless **earth** would reproach

you, for indeed your fame reaches wide heaven,

like that of some blameless king who, reverencing the gods,

reigning among men **great in number** and powerful,

upholds good justice, so that the black earth bears

wheat and barley, and the trees are heavy with fruit,

and the flocks produce young without cease, and the sea provides fish

due to his **good rule**, and the people prosper under him.

Now ask me **other things** in your house

— don't **inquire** of my race and **native land**,

lest you fill my heart even more with sufferings

as I am reminded: . . .

110

115

- a Discuss the poetic artistry of this passage (concentrating on the Greek, not the English translation!). You do not need to know Greek to do this. Give examples of phonetic repetition, merisms, or anything else worth noting. For example, the beginning of line 109 (*hôs te teu*) is phonetically mirrored by *hos te theou(dês)* at the end. The vowel letters *a e i o u* are pronounced as in most European languages (Italian, Spanish, German, Latin, etc.); *â, ê*, etc. are long *a, e*, etc.; the sequences *ai, ei, oi, au, eu, ou, âi, êi* are diphthongs and count as single syllables; all other vowel combinations (e.g. *oo, ee*) make up two separate syllables; *x* represents *ks*; and the combinations *ph, th, kh* represent *p, t, k* with a puff of breath after them (as in the first sound of Eng. *pit, tip, kit*).
- b A common feature of IE poetry is *ring composition*, in which a poet uses a word or words toward the beginning of a section and repeats them or phonetically similar words toward the end, forming a "ring" that binds the section together. This section has a double ring that begins with the words *Odu~~ss~~eus* and *gaian* ('land') at the end of lines 106 and 107. Identify the two forms, and their location, that close this ring. One of these may not be immediately obvious.

- 6 Write a paragraph comparing the content of the middle of the above excerpt from the *Odyssey*, where Odysseus talks about a just ruler, with the following excerpt from the Old Irish *Audacht Morainn* (*Testament of Morann*), a set of instructions for princes (§§17–21, translated by Fergus Kelly; for the original, see §14.50): “It is through the justice of the ruler that abundances of great tree-fruit of the great wood are tasted. It is through the justice of the ruler that milk-yields of great cattle are maintained. It is through the justice of the ruler that there is abundance of every high, tall corn. It is through the justice of the ruler that abundance of fish swim in streams. It is through the justice of the ruler that fair children are well begotten . . .”

3 Proto-Indo-European Phonology

Introduction

3.1. The science of linguistic reconstruction goes back to work of the German philologist August Schleicher in the 1860s. He imputed only three vowels to Proto-Indo-European (*a i u*) and a total of fifteen consonants. A rush of discoveries in the later 1860s and 1870s revealed that this reconstruction was inadequate. By the close of the nineteenth century, the model of PIE phonology that is presented in this chapter, with 25 consonants and 10 vowels, was mostly in place.

The notation of the reconstructed sounds of PIE has its own history. Readers familiar with the International Phonetic Alphabet (IPA), a universal phonetic transcription system devised in the late nineteenth century, need to be aware that the notation of PIE sounds is similar to, but not identical with, that of the IPA. As is standard, whenever recourse is made to IPA symbols, they will be enclosed in square brackets.

Consonants

3.2. Consonants are speech sounds produced by partially or completely blocking the flow of air through the vocal tract. They are classified according to where in the vocal tract this obstruction occurs (the *place* of articulation), how the obstruction is produced (the *manner* of articulation), and whether there is accompanying vibration of the vocal cords (*voicing*).

Stops

3.3. Stop consonants (or plosives) are produced by completely blocking, then releasing, the flow of air through the mouth. From the comparative evidence afforded by the most ancient IE daughter languages, we know that PIE was rich in stop consonants. Over the next couple of pages, we will examine the evidence and briefly show the reasoning behind the standard reconstruction of the PIE stop inventory. (Space will not allow similar treatment of the sounds discussed in the rest of this chapter; the reader may bear in mind that similar reasoning was employed for all of the other reconstructions.)

3.4. Let us begin with the following sound correspondences from a selection of IE daughter languages, a bit simplified for our present purposes:

Sanskrit	Avestan	Greek	Latin	Welsh	English	Armenian	Lithuanian
p	p	p	p	—	f	h	p
t	t	t	t	t	th	tʰ	t
b	b	b	b	b	p	p	b
d	d	d	d	d	t	t	d

(The symbol *tʰ* in Armenian represents a *t* with aspiration, a following puff of breath.) All of these are *reflexes*, that is, descendant outcomes, of ancestral PIE cognates. This table is a shorthand for saying that, for any Sanskrit word containing *p*, its Avestan, Greek, Latin, and Lithuanian cognates will also contain a *p* in the same position, while the English cognate will have *f*, the Armenian cognate will have *h*, and the Welsh cognate will have no sound. (We will provide actual word equations illustrating these sound correspondences below in §3.7.) The same is true, *mutatis mutandis*, for the other sounds listed above, where all the languages agree across the board except English and Armenian. Cross-linguistic research has shown that it is more natural for the English and Armenian inventories to have developed from *p t b d* than the other way around; furthermore, if we hypothesize that only English and Armenian have changed the original state of affairs (and Welsh in the case of *p*), we are not forced to claim that all the other branches underwent major independent but parallel (and cross-linguistically unnatural) change.

For these and other reasons, on the basis of these comparanda Indo-Europeanists reconstruct the voiceless stops **p* and **t* and the corresponding voiced stops **b* and **d* for PIE; they were pronounced as in English. **p* and **b* are both labial stops, and **t* and **d* are dentals. (Precisely what kind of dental stop **t* and **d* were is not certain, as “dental” is a cover term for a variety of sounds produced by the tip of the tongue touching the upper teeth; but this level of phonetic detail is not likely to be recoverable for PIE.)

3.5. A second set of correspondences is more complicated, and caused early Indo-Europeanists considerable trouble:

Sanskrit	Avestan	Greek	Latin	Welsh	English	Armenian	Lithuanian
k	k	k	k¹	c	h	kʰ	k
ś	s	k	k¹	c	h	s	š
k	k	p	qu	p	wh	kʰ	k
g	g	g	g	g	k²	k	g
j	z	g	g	g	k²	c	ž
g	g	b	u³	b	qu	k	g

¹ Spelled *c*. ² Often spelled *ç*. ³ Or *v*; pronounced *w*.

These sets all involve *velar* sounds of various kinds – *k*’s and *g*’s – produced with the back of the tongue raised to touch the soft palate (velum). The big issue was how many velars to reconstruct for PIE. It turns out that the only satisfactory solution is to reconstruct a separate sound for each of these correspondence sets. For the first and fourth rows, all the languages except English and Armenian (where *kʰ* stands for a *k* with aspiration) agree on having ordinary *k* and *g* (Welsh *c*

represents *k*). We therefore reconstruct **k* and **g* for these two correspondence sets. (This **g* is the "hard" *g* as in English *garden*.)

The second and fifth rows include plain velars in some languages but a palatal stop (written *ś*) in Sanskrit, an alveopalatal affricate (*j*, as in English) in Sanskrit, a collection of sibilants (*s*, *z*, *ś*, *ž*, the last two like English *sh* and *zh*), and an alveolar affricate (*c* in Armenian, pronounced *ts*). All these sounds commonly develop out of palatal or palatalized sounds in the histories of many of the world's languages. It is therefore believed that these sounds go back to *palatal* stops in PIE, written **k̑* and **g̑*, pronounced like ordinary *k* and *g* but farther forward in the mouth (as in *Tokyo* and *argue*). In some languages, such as Greek, Latin, and English, the palatal velars fell together with the ordinary velars, more on which later.

This leaves the third and sixth sets above, both of which distinguish themselves by having a labial element of some kind in several of the branches: labial stops (*p b*) in Greek and Welsh, and a labiovelar sound in Latin and English (*qu, wh*) – that is, a velar sound with rounding of the lips. There is essentially only one kind of sound that could have had such diverse developments (as either plain velar or labiovelar): a *labiovelar stop*. We therefore reconstruct two PIE labiovelars, **kʷ* (voiceless; also written **kʰ*) and **gʷ* (voiced; also written **gʰ*). These were single speech sounds, velars pronounced with rounding of the lips; they were not sequences of velar plus *w*. (We know this because such sequences have different outcomes from the labiovelars in some languages. In Sanskrit and Lithuanian, the PIE labiovelar **kʷ* became *k*, but the consonant cluster **kʷ* (= *kw*) became *kv*.)

Although the term *velar* normally refers specifically to *k* or *g*, all the six sounds we have just reconstructed – **k* **g* **k̑* **g̑* **kʷ* **gʷ* – are collectively referred to as "velars" in Indo-European studies (or "gutturals" in older literature; occasionally they are also called *dorsals* or *tectals*). To avoid confusion, **k* and **g* are often called the *plain velars* to distinguish them from the others (which, as we have seen, are called palatal velars and labiovelars).

3.6. There is one final set of correspondences involving stop consonants:

Sanskrit	Avestan	Greek	Latin	Welsh	English	Armenian	Lithuanian
bh	b	ph	f	b	b	b	b
dh	d	th	f	d	d	d	d
h	z	kh	h	g	g	j	ž
gh	g	kh	h	g	g	g	g
gʰ	g	ph	f	g	b	g	g

In three of the branches above, these sounds are characterized by aspiration. Sanskrit *bh* is a *b* followed by a puff of breath; Greek *ph* was a *p* followed by a puff of breath; Latin *f* is aspirated by its very nature. On the other hand, English, Armenian, and some other branches do not show any aspiration here. It is generally agreed that the aspiration must have been originally present, but lost in the branches that no longer show it; the opposite direction of change is unlikely. We reconstruct five so-called *voiced aspirated stops*, symbolized **bʰ* **dʰ* **gʰ* **gh* **gʰʰ*. The voiced aspirates can also be written with superscript *h*'s (*bʰ gʰ* etc.); they were voiced stops followed by a puff of breath or a brief period of breathy voice (the latter technically called *murmur*). Indic (with its main ancient representative Sanskrit) is the only

subbranch of the family that preserved them intact into the historical period; they are still found in many modern languages descended from Sanskrit, such as Hindi and Bengali.

(We may add parenthetically that in earlier Indo-European studies a fourth series of stops was reconstructed, the *voiceless aspirates* **ph* **th* **kh* etc. The best evidence for them comes from Indo-Iranian and Greek, where however they can be explained as later, secondary developments. A few Indo-Europeanists still posit them for the proto-language. We will discuss these a bit more in §10.31.)

To sum up so far, PIE had the following fifteen-stop inventory:

	labial	dental	palatal	velar	labiovelar
voiceless	p	t	k̑	k	kʷ
voiced	b	d	g̑	g	gʷ
voiced aspirated	bh	dh	gh̑	gh	gʷh

3.7. Following are representative cognate words to illustrate each of the correspondence sets so far discussed. If no English gloss (translation) is given for a particular form, it has the same meaning as its PIE ancestor. Not all the possible cognates are given in each set.

The forms cited as examples in this and the next few chapters may seem a bit bewildering at this stage, as we have not yet covered the various spelling systems of these languages, nor the sound changes that account for any deviations from a particular reconstructed PIE form. For the time being, such details are not significant and should be ignored; these lists are designed mainly to give a sense of the large body of evidence that the reconstructions are based on. In later chapters, we will have occasion to explicate the sound changes that caused any such deviations observed below. There are also a few forms in the lists below where an outcome is different from that predicted by the charts in §§3.4–6; these are typically due to a sound change conditioned by a particular neighboring sound and should not cause concern at this point.

As most of the IE language names are not assumed to be familiar at this stage, only the following abbreviations are used: Eng. = English; Gk. = Greek; Lat. = Latin; Vedic Skt. = Vedic Sanskrit (the oldest form of Sanskrit and the one generally cited in IE studies). Hyphens are used to separate prefixes and suffixes from root morphemes wherever this clarifies an example; a form ending in a hyphen means that normally the form had to have a grammatical ending attached to it, or that it was a prefix.

VOICELESS STOPS

These typically remain unchanged, except in Germanic and Armenian, where they became fricatives or aspirates. The labial *p* is sometimes weakened or lost (Celtic, Armenian).

**p* **ped-*, **pod-* ‘foot’: Luvian *pāta-*, Vedic Skt. *pad-*, Gk. *pod-*, Lat. *ped-*, Armenian *botn*, Eng. *foot*, Tocharian A *pe*, Albanian *për-posh* ‘under’ (< **pēdsu* ‘at the feet’)

- **t* **ters-* 'to dry': Vedic Skt. *tarṣ-āyati* 'makes thirsty', Gk. *tērs-etai* 'becomes dry', Lat. *terra* '(dry) land', Eng. *thirst*, Armenian *t'ar-amim* 'I wilt', Albanian *ter* 'I dry'
- **k* **kerd-* 'heart': Hieroglyphic Luvian *zārza*, Hittite *kard-*, Skt. *śrād-dhā-* 'putting one's heart (in), trust', Gk. *kard-iā*, Lat. *cord-*, Old Irish *cride*, Eng. *heart*, Armenian *sirt*, Tocharian B *kār-yāñ* 'hearts', Old Church Slavonic *srīd-ice*, Lithuanian *širdis*
- **ḱ* **kes-* 'comb': Hittite *kišš-*, Gk. *késkeon* (< **kes-kes-o-*) 'coarse flax prepared for spinning, tow', Middle Irish *cír* (< **kēs-rā*) 'comb', Old Eng. *beord* 'hair' (> Modern Eng. *hards* 'coarse refuse of flax'), Old Church Slavonic *kosa* 'hair', Lithuanian *kasà* 'braid'
- **kʷ* **kʷi-*, **kʷo-* (interrogative pronoun stems): Hittite *kuit* 'what', Vedic Skt. *kás* 'who', Avestan *kō* 'who', Gk. *tí* 'what', Lat. *quid* 'what', Welsh *prwy* 'who', Armenian *k'ani* 'how much', Old Eng. *hwæt* 'what', Tocharian B *kʷse-* 'which', Old Church Slavonic *kū-to* 'who', Lithuanian *kàs* 'who', Albanian *kë* 'whom'

VOICED STOPS

These also typically remain unchanged, the main exceptions being Germanic, Armenian, and Tocharian, where they became voiceless:

- **b* **bel-* 'strong, strength': Vedic Skt. *bálam* 'strength', Gk. *bel-tíon* 'better', Lat. *dē-bilis* 'lacking strength', Old Church Slavonic *boljī* 'bigger'
- **d* **doru*, **deru* 'wood, tree': Hittite *tāru*, Vedic Skt. *dāru*, Avestan *dāuru*, Gk. *dóru*, Armenian *tram* 'firm', Old Irish *daur* 'oak', Eng. *tree*, Old Church Slavonic *drěvo*, Lithuanian *dervà*, Albanian *dru* 'wood'
- **ǵ* **ǵonu*, **ǵenu* 'knee': Hittite *ǵēnu*, Vedic Skt. *jānu*, Avestan *žnūm* (accus.), Gk. *gónu*, Lat. *genū*, Armenian *counr*, Eng. *knee*, Tocharian A *kanwem* 'the two knees'
- **g* **gras-* 'eat': Vedic Skt. *grásate* 'eats, feeds', Gk. *grástis* 'green fodder, grass', Lat. *grāmen* (< **gras-men*) 'grass, fodder'
- **gʷ* **gʷen-* 'woman': Hittite *kuinnaš*, Vedic Skt. *jāni-*, Avestan *jaini-*, Gk. *guné*, Old Irish *ben*, Old Eng. *cuēn* (Mod. Eng. *queen*), Armenian *kin*, Tocharian B *šana*, Old Church Slavonic *žena*, Old Prussian *genna*

VOICED ASPIRATED STOPS

These are preserved intact only in Indic (as per §3.6 above). They fell together with the plain voiced stops in most of the other branches, but have reflexes distinct from those of the voiceless and plain voiced stops in Greek, Italic, Germanic, and Armenian (and Indic, of course).

- *bh *bher- 'carry': Vedic Skt. *bhārami* 'I carry', Gk. *phērō* 'I carry', Lat. *ferō* 'I carry', Phrygian *ab-beret* 'he brought', Old Irish *biru* 'I carry', Armenian *berem* 'I carry', Eng. *bear*, Tocharian A and B *pār-*, Old Church Slavonic *berq* 'I take', Albanian *bjer* 'bring!'
- *dh *medhu 'honey; sweet drink': Luvian *maddu* 'wine', Vedic Skt. *mádhu* 'honey', Gk. *méthu* 'wine', Old Irish *mid* 'mead', Eng. *mead*, Tocharian B *mit* 'honey', Old Church Slavonic *medŭ* 'honey', Lithuanian *medūs* 'honey'
- *gh *gheu- 'pour': Vedic Skt. *hūyáte* 'is poured', Avestan *zaotar-* 'priest (who pours libation)', Gk. *khé(w)ō* 'I pour', Lat. *fūtis* 'watering-can', Armenian *joyl* 'having been poured', Tocharian B *kewu* 'I will pour', Eng. *ingot*
- *gh *steigh- 'go, climb': Vedic Skt. *ati-ṣṭigham* 'to climb up', Gk. *steikhō* 'I climb', Old Irish *tíagu* 'I go', Gothic *steigan* 'to climb', Old Church Slavonic *stignōti* 'to come, arrive at', Lithuanian *steigiūos* 'I hurry', Albanian *shteg* 'path'
- *g^wh *g^when- 'smite, slay': Hittite *kuen-zi* 'slays', Vedic Skt. *hán-ti* 'slays', Avestan *jain-ti* 'slays', Gk. *-phonos* '-slayer', Lat. *dē-fen-dit* 'beats off, defends', Old Irish *gon-im* 'I slay', Eng. *bane*, Armenian *jn-em* 'I slay', Old Church Slavonic *ženq* 'I drive after, hunt', Lithuanian *genū* 'I drive'

"CENTUM" VS. "SATEM" DEVELOPMENT OF THE VELARS

3.8. The three series of velar consonants (palatals, velars, and labiovelars) collapsed to two in almost all the IE daughter languages, with either the palatals or the labiovelars merging with the plain velars. In the first case, the result was a plain velar series contrasting with a labiovelar series (*k g gh* and *k^w g^w gh*). In the second case, the result was a plain velar series contrasting with a palatal series (*k g gh* and *k̑ g̑ gh̑*); in branches exhibiting this development, as we saw above, the palatals typically went on to become affricates (*ch-* or *ts-*like sounds) and/or sibilants (*s-* and *sh-*like sounds). As a shorthand for referring to these two directions of development, Indo-Europeanists use the terms *centum* and *satem*, from the Latin and Avestan words for 'hundred' (*centum* and *satəm*, respectively), whose initial sounds come from PIE **k̑*. The Latin word begins with the sound *k* (spelled *c*) and therefore symbolizes the merger of palatal **k̑* with ordinary **k*, while in Avestan, **k̑* stayed distinct from **k* and became the sibilant *s*. The two developments are summarized below:

centum	PIE	satem
*k {	*k̑	*k̑
*k ^w	*k	*k

3.9. For a long time it was thought that the distinction between centum and satem languages reflected an old dialectal division within IE, particularly since the

two groups appeared not to overlap geographically: the centum branches (Greek, Italic, Celtic, Germanic) are more westerly than most satem branches (Indo-Iranian, Armenian, Balto-Slavic). However, discoveries in the twentieth century complicated this view. Tocharian, located farther to the east than any other branch, is centum (if it can be called anything – all three series fell together as one for the most part), and the Anatolian language Luvian appears to have distinct reflexes of all three series (see §9.48). In addition, many words in Balto-Slavic, which is satem, actually show centum developments (see §18.5), and Albanian, which is most likely satem, lies in a generally centum area. It is therefore at least possible that each branch became centum or satem independently, although this view adds complications of its own. In any case, the terms centum and satem remain a useful descriptive shorthand.

3.10. Before leaving the realm of the PIE stop consonants, mention should be made of the glottalic theory, the most prominent alternative theory to the traditional view of the PIE stop system. Its starting point was a famous statement by the linguist Roman Jakobson in 1957 that the reconstructed consonant inventory of PIE was typologically improbable or impossible. This eventually spurred Tamaz V. Gamkrelidze and Vyacheslav V. Ivanov, and independently the American linguist Paul Hopper, to set up a different system in 1973. According to these scholars, PIE possessed a series of glottalized voiceless stops (stops followed by a brief closure of the glottis, symbolized $p' t' k'$ etc.; technically called *ejectives*) in place of the traditional voiced stops $b d g$ etc. They and their followers have averred that the resultant system is typologically more natural than the traditional one. Figuring into their arguments is the rarity of $*b$, which appears in only a handful of reconstructed words; in their view, if we reconstruct glottalized $*p'$ instead of $*b$, then PIE would have been just like some other languages with glottalized voiceless stops but lacking p' . They further claim that the glottalic system accounts for the curious fact (which we will get to in §4.9) that two plain voiced stops cannot co-occur in the same root (thus $*deb-$, $*ged-$, etc. are impossible PIE roots): if in fact the constraint was one against the co-occurrence of two glottalized consonants in the same root ($*t'ep'$, $*k'et'$, etc.), then that would be a familiar constraint known from other languages possessing such consonants.

3.11. The glottalic theory enjoyed a not insignificant following for a time, and still has adherents; but it has been rejected by most Indo-Europeanists. A full discussion cannot be embarked upon here, but the following points may be mentioned. Jakobson's original claim has since been falsified by the discovery of some languages having stop inventories structurally comparable to PIE. A second problem is the complete lack of direct comparative evidence for glottalized stops in PIE. None of the daughter languages has them except some dialects of Armenian and the Iranian language Ossetic, both of which acquired them in recent times from neighboring non-IE languages. This makes accounting for the phonological histories of nearly all the branches much more complicated than under the traditional framework. Finally, the arguments based on the status of $*b$ are weak, for $*b$ is merely rare in PIE, not absent, and the statistical frequency of a sound does not necessarily indicate anything about that sound's history.

Fricatives

3.12. Fricatives are consonants produced with only partial closure of the vocal tract and typically having a hissing or buzzing quality. PIE possessed the sibilant fricative *s, as in the root for 'sit', *sed-: the s is preserved intact in e.g. Vedic Skt. *sáda* 'sit!', Lat. *sedere* 'to sit', Eng. *sit*, and Old Church Slavonic *seděti* 'to sit'. In some other languages, notably Greek, it became h: Gk. *hézomai* 'I sit down', Armenian *hecanim* 'I sit, ride'.

PIE *s changed phonetically to *z before voiced stops (by the same rule discussed below in §3.34): a form of the root for 'sit' is found in the compound *ni-sd-o- 'where (the bird) sits down = nest', which was pronounced *nizdo-, with the z preserved intact in Old Church Slavonic *gnězdo* 'nest' (with later added g-) and Lithuanian *lizdas* 'nest' (with later replacement of the initial n- by l-). In Vedic Skt. *nīdás* and Lat. *nīdus*, this *z has disappeared but lengthened the preceding vowel.

Resonants: Liquids, nasals, and glides

3.13. PIE had six resonants, consonants produced with little obstruction of the flow of air and having a sonorous quality: the liquids *l and *r, the nasals *m and *n, and the glides (or semivowels) *i and *u (also written as in English, *y *w). Technically the nasals are stops, but since air flows freely during their pronunciation (namely, through the nose), they behave like the liquids and glides and are always considered together with them in IE studies.

The liquids and nasals are preserved intact in most of the branches:

- *l *leuk- 'light': Hittite *lukke-* 'kindle', Vedic Skt. *rócate* 'shines', Avestan *raoc-aiēsti* 'causes to shine, lights up', Gk. *leukós* 'bright', Lat. *lūc-* 'light', Old Irish *luchair* 'a shining', Armenian *loys* 'light', Eng. *light*, Tocharian A and B *luk-* 'to shine', Old Church Slavonic *luča* 'beam of light'
- *r *pro, *prō 'forward': Hittite *parā* 'forth', Vedic Skt. *prá*, Avestan *fra-*, Gk. *pró*, Lat. *prō* 'in front of', Old Irish *ro-* (perfective and intensive prefix on verbs), Eng. *fro*, Old Church Slavonic *pro-* 'through', Lithuanian *prā* 'past'
- *m *men- 'think': Vedic Skt. *mānas-* 'mind', Avestan *manah-* 'mind', Gk. *ménos* 'mental energy, spirit, wrath', Lat. *mēns* 'mind', Old Irish *do-moiniur* 'I think', Armenian *i-manam* 'I understand', Eng. *mind*, Old Church Slavonic *mīnjō* 'I believe', Lithuanian *menù* 'I think'
- *n *ne 'not': Hittite *na-tta* 'not', Vedic Skt. *ná*, Avestan *na*, Lat. *ne-* (as in *ne-scīre* 'not to know'), Old Irish *ní* (< *nē), Old Eng. *ne*, Old Church Slavonic *ne*, Lithuanian *nè*

The glides underwent change more frequently than the other resonants. The palatal glide *i sometimes weakened and disappeared, and the labiovelar glide *u typically became a fricative v or f. Interestingly, only in West Germanic (which includes English) has PIE *u survived unchanged to the present day, as w.

- *j *iugom 'a yoke': Hittite *iukan* (phonetically *yugan*), Vedic Skt. *yugám*, Gk. *zugón*, Lat. *iugum*, Welsh *iau*, Eng. *yoke*, Lithuanian *jūngas* (with added n)

- **u* **ueǵh*- 'to lead, convey (in a vehicle)': Hieroglyphic Luvian *waza*- 'drive', Vedic Skt. *váhati* 'leads, brings', Avestan *vazaiti* 'leads, brings', Gk. (Pamphylian) *wekhetō* 'let him convey', Lat. *uehere* 'to convey', Old Irish *fén* (< **ueǵh-no*-) 'wagon', Middle Dutch *wagen* 'wagon' (borrowed into Eng.), Old Church Slavonic *vezō* 'I convey', Lithuanian *vežū* 'I drive', Albanian *vjedh* 'I steal' (< *'lead away')

SYLLABIC RESONANTS

3.14. When a resonant occurred between two non-syllabic consonants (CRC) or between a consonant and a word-boundary (-CR or RC-), it became syllabic and functioned as a vowel. In IE studies, the syllabic liquids and nasals are conventionally represented with a little circle underneath (**l̥* **r̥* **m̥* **n̥*). They were pronounced much like the liquids and nasals in the second syllable of the English words *bottled*, *battered*, *bottoms*, *buttons* as pronounced in colloquial speech.

In most of the daughter languages, the syllabic resonants **l̥* **r̥* **m̥* **n̥* developed prop vowels before them (such as Germanic **ul* **ur* **um* **un*). Sometimes the syllabic nasals lost their nasality and became plain vowels (as in Greek and Indo-Iranian, where **m̥* and **n̥* became *a*). Below are some representative correspondence sets:

- **l̥* **u**lk̥**ʷos* 'wolf': Hittite *walkuwa*- 'monster' (?), Vedic Skt. *vṛkas*, Avestan *vahrka*-, Old Church Slavonic *vŭlkŭ*, Lithuanian *vilkas*
- **r̥* **m̥**r̥*-to- 'dead': Vedic Skt. *mṛtá*-, Avestan *marata*-, Gk. *brotós* (< **mrotós*) 'mortal', Lat. *Morta* 'goddess of death', Armenian *mard* 'man', Eng. *murd-er*, Old Russian *mŭrtvŭ* 'dead', Lithuanian *mirtis* 'death'
- **m̥* **dek̥m̥* 'ten': Vedic Skt. *dāsa*, Avestan *dasa*, Gk. *déka*, Lat. *decem*, Old Church Slavonic *desę-tŭ* (*ę* is a nasalized vowel), Lithuanian *dėšimt*
- **n̥* **n̥*- 'not, un-': Vedic Skt. *a(n̥)*-, Gk. *a(n̥)*-, Lat. *in-* (older *en-*), Old Irish *an-*, Eng. *un-*

The glides **i* and **u* also had syllabic variants – the vowels **i* and **u*. For comparative examples of their reflexes, see §3.26 below on vowels.

3.15. Earlier handbooks also reconstructed a series of long syllabic resonants, **l̥̄* **r̥̄* **m̥̄* **n̥̄*. These are nowadays understood to have been sequences of syllabic resonant plus laryngeal (see directly below): **l̥̄H* **r̥̄H* **m̥̄H* **n̥̄H*. But the older notation can still be useful in some contexts.

Laryngeals

3.16. The so-called "laryngeals" are a class of sounds whose precise phonetic values are still not known. They are generally agreed to have been fricatives produced in the back of the mouth and throat, and are partly preserved as such in the oldest attested branch, Anatolian (e.g. as the velar fricative written *h* in Hittite and Luvian). Outside Anatolian they have no certain consonantal reflexes, becoming vowels under certain conditions and leaving other, indirect traces. Their discovery by Ferdinand

de Saussure represents one of the triumphs of the comparative method; we will sketch his arguments in §§4.18–19.

In the standard theory, PIE had three of these sounds, represented abstractly as $*h_1$ $*h_2$ $*h_3$ (also written $*H_1$ $*H_2$ $*H_3$ or $*a_1$ $*a_2$ $*a_3$, among other notations). The cover symbol H (less commonly a) is used for a laryngeal that cannot or need not be specified.

3.17. When a laryngeal occurred between two non-syllabic sounds, it developed a prop vowel next to it (sometimes before, sometimes after), yielding what is often called a vocalized or syllabic laryngeal. Like syllabic resonants, vocalized laryngeals may be represented with circles under them: h_1 h_2 h_3 (cover symbol \bar{H}), though we will not be doing that in this book except when necessary. Vocalized laryngeals became vowels in many of the branches, usually a , but i in Sanskrit. In Greek, each vocalized laryngeal had a different outcome, the valuable “triple reflex” $e a o$ from $*h_1$ $*h_2$ $*h_3$ respectively. Some examples showing reflexes of vocalized laryngeals are as follows:

$*h_1$ $*dh_1s-$ ‘sacred, religious’: Vedic Skt. *dhīṣ-nya-* ‘devout’, Gk. *thés-phatos* ‘decreed by god’, Lat. *fānum* (< $*fas-no-$ < $*dh_1s-no-$) ‘temple’

$*h_2$ $*sth_2to-$ ‘standing, made to stand’: Vedic Skt. *sthitá-*, Gk. *statós*, Lat. *status*, Welsh (*gwa*)*stad* ‘standing firm, constant, level’, Old Norse *staðr* ‘obstinate, restive’ (said of horses)

$*h_3$ $*dh_3ti-$ ‘gift’: Vedic Skt. *díti-*, Gk. *dósis*, Lat. *datiō*

In Greek, Armenian, and Phrygian, laryngeals occurring at the beginning of a word before a consonant were also vocalized. Thus $*h_2nēr$ ‘man’ became Gk. *anér*, Phrygian *anar*, and Armenian *ayr* (from earlier $*anir$). In the other branches, the laryngeals did not vocalize in this position and eventually disappeared, though they sometimes left indirect traces that will be discussed in the next chapter.

BEHAVIOR AND EFFECTS OF THE NON-SYLLABIC LARYNGEALS

3.18. There are three important sound changes undergone or induced by the laryngeals when they were not vocalized. The first of these happened already in PIE itself and is known as laryngeal coloring. This refers to the effects that two of the laryngeals, $*h_2$ and $*h_3$, induced on an adjacent short $*e$: when next to $*h_2$, $*e$ became “colored” to $*a$, and when next to $*h_3$, $*e$ became $*o$. Thus $*h_2ent-$ ‘front, forehead’ became colored to $*h_2ant-$, and $*h_3erbh-$ ‘move from one social sphere into another’ became colored to $*h_3orbh-$. The first laryngeal, $*h_1$, did not color; thus the e in $*h_1esti$ ‘is’ remained e .

The coloring effects of laryngeals are what gave them their name; they were dubbed “laryngeals” by Hermann Möller in 1911, inspired by an identically named class of sounds in Semitic languages with similar effects on neighboring vowels. When writing reconstructed forms where coloring could take place, some Indo-Europeanists write them with coloring, some without. In this book, we shall ordinarily write them without coloring (thus $*h_2ent-$ and $*h_3erbh-$), as that usually makes word structures clearer (as we will see in the next chapter).

3.19. The second important development of the non-syllabic laryngeals was their loss when adjacent to a vowel. This process may have begun in PIE, but it did not run its full course until much later. Anatolian, importantly, still preserves consonantal reflexes of $*h_2$ and (most likely) $*h_3$ in many positions as a sound transliterated *h*. Thus the Hittite outcomes of the forms $*h_2ant-$ and $*h_3orbh-$ mentioned above are *hant-* 'forehead' and *harapp-* 'become separated' (the *o*-coloring effect of the third laryngeal is not directly visible in the second example because Hittite changed $*o$ to *a*). The oldest Indo-Iranian languages, Vedic Sanskrit and Old Avestan, also reflect consonantal laryngeals in certain positions as hiatuses or glottal stops. But ultimately, all the consonantal laryngeals were lost in all the non-Anatolian branches. Thus $*h_2ant-$ and $*h_3orbh-$ became *ant-* and *orb-* in Latin *ante* 'before, in front' and *orbis* 'orphan'. As for the first laryngeal, it was lost even in Anatolian: thus $*h_1esti$ 'is' became Hittite *ēšzi*, Gk. *estí*, Lat. *est*.

3.20. If a laryngeal came after a vowel and either stood at the end of a word (-VH) or before another consonant (VHC), the loss of the laryngeal was accompanied by compensatory lengthening of the preceding vowel, yielding $*\bar{V}C$ or $*\bar{V}$. (Another term for this process is *contraction*.) Thus $*dheh_1-m̥$ 'something placed' was contracted to $*dhē-m̥$ (> Gk. (*aná*)-*thēma* 'offering'); $*peh_2-s-$ 'protect' was first colored to $*pah_2-s-$ and then was contracted to $*pā-s-$ (> Lat. *pās-tor* 'shepherd'); $*deh_3-rom$ 'gift' was first colored to $*doh_3-rom$ and then was contracted to $*dō-rom$ (> Gk. *dōron*); and $*pih_1-uerih_2$ 'fat' (feminine) underwent no coloring (since laryngeals did not color $*i$) but did undergo contraction to $*pī-uerī$ (> Vedic Skt. *pīvarī*). As a parallel to this development, compare the change from Old Eng. *niht* to Modern Eng. *night*, where the *-h-* has been lost in pronunciation and the preceding vowel lengthened (and later diphthongized). In Anatolian, as one might expect, compensatory lengthening did not happen where laryngeals were not lost; for example, PIE $*peh_2-s-$ 'protect' above became Hittite *pahš-*.

3.21. No other subject in IE phonology has been as contentious as the laryngeal theory. Resistance to accepting laryngeals altogether (partly because of their somewhat forbidding and algebraic look) has dwindled over time; the combined phonological and morphological evidence for their existence is irrefutable. A lot of Indo-European comparative linguistics is simply not possible without them, for they help to explain too many disparate phenomena too neatly to be dismissed. There are differences of opinion as to how many laryngeals to reconstruct for PIE, however; alternative minority views operate with anywhere from one to a dozen of them, or more. Some maintain that the late Indo-European that developed after Anatolian split off had only one laryngeal. But the triple reflex in Greek makes this position very difficult to maintain.

As regards their phonetic character, one fairly widespread view has it that $*h_1$ was a simple *h* or a glottal stop [ʔ], $*h_2$ a voiceless pharyngeal fricative [ħ] (same sound as the Arabic letter *hā'*), and $*h_3$ a voiced pharyngeal fricative [ʕ] (same sound as Arabic *ʿayn*); but there are other possibilities as well.

Consonant clusters

3.22. PIE was rich not only in consonants but also in consonant clusters, especially at the beginnings of words. Most languages that have clusters arrange the consonants in order of increasing sonority, and this was true of PIE also. Stops and fricatives are the least sonorous consonants, and generally come first, followed by resonants, which are more sonorous. The maximum number of consonants in a

word-initial cluster appears to have been three; longer clusters are sometimes found word-internally.

3.23. The most common word-initial clusters are those where the second consonant is a resonant. Almost all the theoretically possible combinations are found: **pro* 'forward', **dlegh-* 'engage oneself', **k^wrei-* 'buy', **dh_uer-* 'door', **tjeg^w-* 'revere', **g^wieh₂-* 'live', **ġneh₂-* 'know', **pneu-* 'breathe', **sjuH-* 'sew', **sreu-* 'flow', **mreghu-* 'short', **mleuh₂-* 'speak', **uleik^w-* 'flowing, liquid', **ureh₂d-* 'branch, root', **mneh₂-* 'be mindful', **h₂ner-* 'man', **h₃meigh-* 'urinate', **h₃reġ-* 'stretch out the hands', **h₁leudh-* 'go', **h₁ieh₁-* 'throw', **h₂ues-* 'spend the night'. Clusters in which the second consonant is a stop or laryngeal are rare unless the first sound is *s*: **sk_{eh}i-* 'shadow', **spend-* 'pour a libation', **sh_{eh}i₁-* 'bind', **tkei-* 'settle' (see §3.25 below), **pter-* 'wing', **h₃bhel-* 'be of use', **h₃sih₁-* 'be' (optative pl. stem). The two liquids **l* and **r*, the glide **j*, and the nasal **n* did not begin consonant clusters word-initially.

Word-initial clusters with three consonants are also found; examples include **streng-* 'squeeze, tie', **splei-* 'splur', **h₃sieh₁-* 'be' (optative sing. stem), **h₂ster-* 'star', and the imitative root **pster-* 'sneeze'.

3.24. Word-final consonant clusters typically arose through suffixation, as in **euġhest* (= **e-uġh-s-t*) 'conveyed (3rd sing. aorist)' and **bheronts* (= **bher-ont-s*) 'carrying (nominative sing. masc. present participle)'. A number of these were subject to various rules of assimilation (compare §§3.34–35), but we do not really know how a form like **euġhest* was pronounced.

"THORN" CLUSTERS

3.25. The Indo-European languages exhibit a curious set of sound correspondences where Vedic Sanskrit has a cluster *kṣ* (pronounced *ksh*, [kʃ]) corresponding to various clusters in Greek and often single sounds in the other languages. A representative sample of these correspondences is given below; the words mean 'bear', 'earth, ground', and 'waste away, disappear'.

Hittite	Vedic	Avestan	Greek	Latin	Tocharian A
<i>ḫartaggaš¹</i>	<i>ḫkṣas</i>		<i>árktos</i>	<i>ursus</i>	
<i>in-zagan²</i>	<i>kṣam-</i> <i>kṣi-</i>	<i>zam-</i>	<i>khthon-</i> <i>phthi-</i>	<i>hom-ō³</i> <i>si-tis⁴</i>	<i>tkam</i>

¹ Phonetically *ḫartkaš*. ² Cuneiform Luvian, '(putting) in the earth = inhumation'; see further below. ³ 'Person' (originally 'earthling'). ⁴ 'Thirst'.

Under the usual modern view, these correspondences descend from PIE clusters consisting of a dental followed by a velar. The order dental-velar is preserved only in Anatolian and Tocharian. The reconstructions of the forms above are **h₂rk^wtos*, **dhġhom-*, and **dhg^whi-* (the choice of the specific dental and velar in each case is based on other comparative evidence). In the other branches, the two sounds metathesized (switched places), and the dental came to be pronounced as [θ], or its voiced equivalent [ð] if the cluster was a voiced cluster. (In some phonetic alphabets, the sound [θ] is represented by the symbol *þ*, called "thorn", hence the term "thorn cluster.") In many branches, the cluster was simplified to a single consonant. Under another view, though, the dentals in these clusters were pronounced [ts] or [dz],

and were pronounced this way already in PIE before any metathesis had taken place. Recently, the American Indo-Europeanist and Anatolianist Craig Melchert has interpreted Cuneiform Luwian *inzagan-* ‘inhumation’ as preserving precisely this stage: the word was probably pronounced [indzgan] and comes from the phrase **en dhǵhōm* ‘in the earth’.

The behavior of other consonant clusters is treated in §§3.34ff. below.

Vowels

3.26. PIE possessed five short vowels **i *e *a *o *u* (approximately equivalent to the first vowels of *bitter*, *better*, *groggy*, *older*, *super*). The vowel correspondences among the daughter languages present a very confusing picture, and it took over half a century for the relationships to be fully worked out. Several branches merged at least two of the vowels (typically *a* and *o*, as in Indo-Iranian, Germanic, and Balto-Slavic), and all the vowels were subjected to many sound changes in different phonetic environments. The following are examples of reconstructed words with each vowel:

- *i* **mizdho-* ‘reward’: Avestan *mizda-*, Gk. *misthós* ‘wage, pay’, Gothic *mizdo*, Old Church Slavonic *mizda*
- *e* **h₁esti* ‘is’: Hittite *ēšzi*, Vedic Skt. *ásti*, Gk. *estí*, Lat. *est*, Old Irish *is*, Eng. *is*, Old Church Slavonic *jesti*, Lithuanian *ėsti*
- *a* **sal-* ‘salt’: Vedic Skt. *sal-ilā-* ‘salty’, Gk. *hál-*, Lat. *sal-*, Armenian *ał*, Old Irish *sal-ann*, Eng. *salt*, Tocharian A *sāle*, Old Church Slavonic *solī*, Lithuanian (dialectal) *salis* ‘sweetness’
- *o* **g^wou-* ‘cow’: Lycian *wawa-*, Skt. *gāv-*, Gk. *bó(w)es* ‘cows’, Lat. *bouēs* ‘cows’, Old Irish *boin* ‘for a cow’, Armenian *kov*, Old Church Slavonic *gov-ęždī* ‘bovine’
- *u* **snusos* ‘daughter-in-law’: Vedic Skt. *snuśā*, Gk. *nuós*, Lat. *nurus*, Armenian *nu*, Old High German *snur*, Old Church Slavonic *snūxa*, Albanian *nuse* ‘bride’

3.27. PIE also had the long non-high vowels **ē *ā *ō*, pronounced similar to the short vowels but held for twice the length. The long vowels **ē* and **ō* are found especially in lengthened-grade forms (see §4.13):

- *ē* **h₃rēǵs* ‘king’: Vedic Skt. *rāṭ*, Lat. *rēx*, Gaulish *-rix* (in personal names)
- *ā* **māter-* ‘mother’: Vedic Skt. *mātār-*, Gk. (Doric) *mātēr*, Lat. *māter*, Armenian *mayr*, Old Irish *máthair*, Old Eng. *mōdor*, Old Church Slavonic *mater-*, Lithuanian *mótė*, Albanian *motër* ‘sister’ (originally ‘older sister taking on the mother’s role’?)
- *ō* **suesōr* ‘sister’: Vedic Skt. *svāsā*, Lat. *sorōr-*, Old Irish *siur*

Whether PIE had **ī* and **ū* is not fully certain because so many, if not all, of the putative examples may have come originally from **iH* or **uH*, which had not

contracted to **ī* and **ū* until after the breakup of the family. (This is true of many of the traditional examples of **ē*, **ā*, and **ō* as well.) However, the long *ū* in the adverbs **nū* and **nūn* 'now' (as in Gk. *nū(n)*, Vedic Skt. *nūn-ām* 'now', Lat. *nū-per* 'recently') is almost certainly not from **uH*, and thus represents a genuine old PIE **u*. Other words, such as **mūs* 'mouse' (in e.g. Gk. *mūs*, Lat. *mūs*, Old Eng. *mūs*) and **dhghū-* 'fish' (in e.g. Gk. *ikhthūs*, Armenian *jow-kn*), may have contained a laryngeal. Also uncertain is the word **uīs* 'poison' (as in Avestan *viš*, Lat. *utr-us*, Middle Irish *fí*), which had a short-vowel variant **uis* (e.g. Vedic Skt. *viṣ-* 'excrement'); but similar words having long- and short-vowel variants are usually thought to have contained a laryngeal that disappeared under unknown conditions in PIE without compensatory lengthening. However, for the purposes of the phonological history of most of the daughter languages, we need not worry about these issues and can reconstruct **ī* and *ū*.

Some long vowels arose by contraction of two short vowels, especially in certain inflectional categories such as the thematic subjunctive (§5.51). Dispute surrounds most or all of the remaining long vowels: since the laryngeal theory demonstrated that long vowels could arise from an earlier sequence of short vowel plus laryngeal, some researchers have claimed that all PIE long vowels were once short vowels followed by laryngeal. But this is suspect without clear evidence underpinning it. Rewriting the traditional reconstructions **māter-* 'mother' and **bhrāter-* 'brother' as **meh₂ter-* and **bhreh₂ter-* really gains us nothing, as it provides no fresh insight into either the morphological or the semantic makeup of these words.

Diphthongs

3.28. The non-high vowels **e* **a* **o* could combine with a following high vowel **i* or **u* to form six diphthongs: **ei* **ai* **oi* **eu* **au* **ou*. Examples include **kei-* 'to lie' (in e.g. Gk. *keimai* 'I lie'), **kaiko-* 'one-eyed' (in e.g. Lat. *caecus* 'blind', Gothic *haihs* 'one-eyed'), **moino-* 'exchange, reciprocation' (in e.g. Archaic Lat. *moenus* 'duty, tribute, payment'), **sreu-mṇ* 'a flow' (in e.g. Gk. *rheūma*), **sauso-* 'dry' (in e.g. Gk. *haños* and Lithuanian *sausas* 'dry'), and **h₂kous-* 'hear' (in e.g. Gk. *akoúō* 'I hear'). Because the high vowel in such diphthongs is rather glide-like in pronunciation, it is common to see these diphthongs written **ei*, **eu*, etc. This is not done in the present work.

Long diphthongs, composed of a long vowel plus **i* or **u*, are rare. In PIE they typically arose by contraction or due to morphologically induced lengthening of a root vowel (the so-called lengthened-grade, discussed in the next chapter). In later dialectal PIE, various long diphthongs were created by contraction processes involving laryngeals. By the time of the attested histories of the daughter languages, however, the long diphthongs had usually been shortened or monophthongized. Examples include **dēik-s-* 'showed' (aorist tense, in Avestan *dāiš* 'he showed'), the dative singular ending **-ōi* (in e.g. Gk. *hipp-ōi* 'for a horse'), **diēus* 'sky, sky-god' (in e.g. Vedic Skt. *dyāus* 'sky-god'; shortened in Gk. *Zeús* 'Zeus'), and **g^wōus* 'cow' (in e.g. Avestan *gāuš*).

"Schwa secundum"

3.29. In word-initial consonant clusters consisting of two stops plus a resonant, a prop vowel was introduced between the two stops to break up the cluster; this vowel is called "schwa secundum" (as opposed to the "schwa primum," the schwa

that represents a laryngeal, §3.16). It is written with a little subscript ̥ or schwa. For example, one of the forms of the numeral 'four' was $*k^w\text{̥}tuor-$, which developed a schwa secundum to become $*k^w\text{̥}tuor-$ or $*k^w\text{̥}tuor-$, yielding Lat. *quattuor* and Homeric Gk. *písures* (as opposed to Gk. *téssares* from a different form, $*k^w\text{̥}etuor-$).

Stress and accent

3.30. Most polysyllabic words have at least one syllable that is acoustically more prominent than the others; this is the stressed or accented syllable. Linguists distinguish broadly between stress-accent languages, where the primary feature of this syllable is greater volume (as in English or German), and pitch-accent languages, where the primary feature of this syllable is a different pitch or pitch contour from that of its neighbors (as in Ancient Greek or Japanese).

From the available comparative evidence, it is standardly agreed that PIE was a pitch-accent language. There are numerous indications that the accented syllable was higher in pitch than the surrounding syllables. Among the IE daughters, a pitch-accent system is found in Vedic Sanskrit, Ancient Greek, the Baltic languages, and some South Slavic languages, although none of these preserves the original system intact.

3.31. All full lexical content words in PIE (nouns, verbs, adjectives, and most adverbs) had an accented syllable. Finite verbs and the vocatives of nouns had idiosyncratic stress properties; see §§5.59 and 6.7. A certain number of function words (some particles, pronouns, and conjunctions) had no stress and were clitics, that is, they could not stand alone and were phonologically attached to a neighboring stressed word with which they formed a single accentual unit. (Compare the *'em* in English *I see 'em*, which cannot stand on its own and is pronounced together with the preceding word *see*.)

3.32. An additional feature of the accent in PIE was that it was mobile, that is, its position was not predictable on the basis of any phonetic features of the word it appeared in. A mobile accentual system is continued in Hittite, Vedic Sanskrit, Greek, Tocharian, and parts of Balto-Slavic, but the other branches have developed fixed accent systems or systems in which accent position is predictable for other reasons. In mobile accent systems, words can be distinguished by the position of the accent alone: Vedic Skt. *úṣas* 'o dawn' (vocative sing.) vs. *uṣás* 'of the dawn'; Gk. *oĩkoi* 'at home' vs. *oĩkoi* (with ̅ indicating accentuation of the first rather than second half of the diphthong) 'houses'; Russian *pisál* 'he was writing' vs. *písal* 'he was peeing'.

In this book, the accent will normally not be shown on reconstructed PIE forms unless it is important in context. Much of the time, in fact, we do not know on which syllable the accent fell.

Phonological Rules

3.33. As with all living languages, the pronunciation of the sounds of PIE varied depending on what other sounds they were combined with in any given word or phrase. This behavior can be captured in statements called phonological rules. Evidence for some of these rules is preserved in several of the daughter languages.

Except for certain basic and near-universal principles of syllabification, all phonological rules are the result of sound change. Some of the rules below represent

features that were part of the living grammar of PIE, while others are fossilized remnants of older sound changes. It is not always possible to tell, for any given rule, which one we are dealing with. Thus the rules listed below may well belong to more than one chronological layer of the proto-language.

Rules of consonant clusters

VOICING ASSIMILATION

3.34. In a cluster of two consonants differing in voicing, the voicing of the first assimilated to (that is, became like) that of the second. Most frequently, the second consonant in such clusters is voiceless, causing the preceding consonant to become voiceless as well: **leḡ-tó-* 'chosen' became **lek-tó-* (Gk. *lektós*), and **h₁éd-si* 'you (sing.) eat' became **h₁ét-si* (Hittite *ēzzi* [pronounced *étsi*], Vedic Skt. *átsi*). The reverse development is seen in **ni-sd-o-* > **ni-zd-o-* 'nest' (cp. §3.12 above), as well as in Avestan *fra-bda-* 'front of the foot' from **pro-pd-o-* (from the root **ped-* 'foot'). Assimilation of this sort, from right to left (the voicing feature can be thought of as spreading leftward), is called *regressive* or *anticipatory assimilation*.

3.35. Assimilation in the reverse direction, called *progressive* or *perseverative assimilation*, is found in one famous example known as Bartholomae's Law. This is reflected most clearly in Indo-Iranian; whether it was of PIE date is controversial. We will discuss it in §10.6.

DENTAL-PLUS-DENTAL CLUSTERS

3.36. A sequence of two dental consonants was pronounced with an added sibilant inserted between them. Thus **-tt-* (from either **-tt-* or **-dt-*, by voicing assimilation) and **-dd-* were pronounced **-tst-* and **-dzd-*. The first was very common, as it happened anytime a root-final *t* or *d* was followed by a suffix beginning with *t* (of which there were many). This cluster survives as *tst* in Anatolian, but in the other branches it was simplified to *-tt-* (Indic), *-st-* (Iranian, Greek, Balto-Slavic), or *-ss-* (Italic, Celtic, Germanic). Thus **h₁ed-te(-)* 'eat! (pl.)' first became **h₁ette(-)* by voicing assimilation, and then **h₁etste(-)*, preserved in Old Hittite *ēzten* (pronounced *étsten*) but simplified in Vedic Skt. *attá*. Similarly, **h₁éd-ti* 'eats' > Lithuanian *ėsti* and Old Church Slavonic *jastŭ*; **uid-to-* 'known' > **uitto-* > **uitsto-* > Gk. (*u*)*istós*, Old Irish *-fess*, and Germ. *ge-wiss* 'certain'.

SIMPLIFICATION OF **ss*

3.37. The cluster **ss*, arising at morpheme boundaries where a suffix beginning with an *s* was added to a root or stem ending in the same sound, was simplified to single **s*. Thus the 2nd singular verb **h₁es-si* 'you are' was simplified to **h₁esi* already in PIE, as shown by Vedic Skt. *ási* and Greek *eī* (< **ehi* < **est*).

STANG'S AND SZEMERÉNYI'S LAWS

3.38. The Indo-Europeanist and Balticist Christian Stang noticed that certain word-final consonant clusters containing resonants appear to have simplified, with compensatory lengthening of a preceding short vowel. The best examples include the

accusative singulars **djēm* 'sky' and **g^wom* 'cow' from earlier **djēum* and **g^woum* (yielding Vedic Skt. *dyām* and *gām*, and Homeric Gk. *Zēn* 'Zeus' and *bōn*). It also operated when the consonant preceding the *-m* was a laryngeal, as witnessed by the feminine accusative singular in **-eh₂m* (see §6.70), which became **-ām* instead of disyllabic **-ah₂m* as expected by the ordinary rules of syllabification in PIE (see §3.42 below).

Similar to Stang's Law is Szemerényi's Law, named after the Hungarian linguist Oswald Szemerényi. It states that **-VRs* (where *R* is a resonant) became **-VR̥*. Thus an earlier **ph₂ters* 'father' and **k₂yons* 'dog' became **ph₂tēr* and **k₂yōn*. The law was devised to explain morphological peculiarities of these words; see further §6.6.

THE BOUKÓLOS RULE

3.39. A labiovelar lost its labial element when adjacent to the vowel **u*. This is an example of *dissimilation*, whereby two sounds that are alike in some way become less alike (both labiovelars and the vowel *u* are pronounced with rounding of the lips). The classic example is Greek *bou-kólos* 'cowherd' from **g^wou-kolos*, dissimilated by the rule from **g^wou-k^wolos*. If the labiovelar had been preserved and not undergone dissimilation, the word should have turned out as **bou-pólos* in Greek; compare the related term *ai-pólos* 'goatherd' < **ai(ǵ)-k^wolos*.

OTHER RULES

3.40. Consonants at the ends of words were subject to special rules. Voicing distinctions in word-final stops were neutralized, so for example word-final *-t* and *-d* were pronounced the same. (Whether they were pronounced voiced or voiceless is a bit unclear.) Additionally, we know that word-final **-n* was lost after *ō*. Contrast Old Irish *cú* 'dog' (< **k₂yō* < **k₂yōn*) and Lat. *homō* 'man' (< **(dh)ǵhemō* < **(dh)ǵhemōn*) with Gk. *humēn* 'membrane' (< **sumēn*) and Old Church Slavonic *ime* 'name' (< **nm-ēn*; the *-ē* is a nasalized vowel). PIE forms where the **-n* disappeared, like the word for 'dog' above, are written by Indo-Europeanists variously with or without the *n*, or with *n* in parentheses: **k₂yōn* or **k₂yō* or **k₂yō(n)*.

3.41. Osthoff's Law. The Indo-Europeanist Hermann Osthoff observed that in many IE languages, long vowels often became short before a resonant plus another consonant. Thus Gothic *winds* 'wind' (earlier **wend-*) and Lat. *uentus* 'wind' both have short vowels, although etymologically they come from **uēnto-* (from even earlier **h₂ueh₂nto-*). Similarly, Gk. *Zeús* 'Zeus' has shortened its vowel vis-à-vis its ancestor, PIE **djēus*. Many of the examples of Osthoff's Law can be explained analogically, and it was not a general rule of PIE date (there is no evidence of it in Indo-Iranian or Tocharian). Pending further research, Osthoff's Law is best regarded as a tendency; such shortening is fairly common cross-linguistically, and the IE examples may have each arisen independently.

Syllabification

3.42. PIE grouped sounds into syllables in much the same way as Greek, Latin, Sanskrit, and many other languages. In any given sequence of consonants and vowels, the vowels constituted the syllabic peaks, and were linked to a preceding consonant (if one was available) which formed the onset (beginning) of a syllable. If two or

more consonants occurred together in the middle of a word, they were usually split between two syllables. In the abstract, a word of the structure VCCVCVCCVC would have been syllabified VC.CV.CVC.CVC. It is possible that certain consonant clusters could group together in the middle of a word as the onset of a syllable; if so, by a universal phonological principle they would have also been able to form word-initial onsets.

Syllabification proceeded from right to left. Let us take a concrete example this time, the word **bherontsu* 'in the ones carrying'. First the vowels – the syllabic peaks – are located, and linked to a preceding consonant: *bhe ro su*. The remaining consonants are then attached to the preceding vowel: *bbe.ront.su*.

One problem currently under discussion concerns the precise rules for syllabifying an unbroken sequence of resonants (with or without laryngeals). It is usually assumed that as one proceeds from right to left, the first resonant or laryngeal in the sequence, as long as it is not followed by a vowel, is made syllabic, the second non-syllabic, and so on. (If the first resonant or laryngeal is followed by a vowel, it is made non-syllabic, while the next one to its left becomes syllabic, and so on.) This rule works for a word like **kynbhi* 'with dogs', but not for a word like **iungenti* 'they yoke' (which should come out as **iungenti* by the usual rule). In the latter case, morphological analogy surely played a role, since the root that this form comes from is **jeug-* or **jug-*, and **jung-* is closer to the root in shape and sound than **iung-*.

Another problem is posed by the feminine suffix **-ih₂* (see §6.71). This became *-ī* in Indo-Iranian and some other branches, but **-ja* in Greek and Tocharian, reflecting different syllabification. The source of the difference is not fully understood.

SIEVERS' AND LINDEMAN'S LAWS

3.43. The Germanic philologist Eduard Sievers noticed that the weight of a syllable in PIE affected the pronunciation of following consonant clusters consisting of a consonant plus a glide before a vowel (e.g. **-tjo-*, **-tuo-*): if the syllable before the cluster was heavy (i.e. if it ended in *-VCC* or *-VC*), the vocalic equivalent of the glide was inserted into the cluster (yielding **-tijo-*, **-tuo-*). Thus a pre-Germanic **bar-jas* 'army' became Gothic *harjis*, but **herdjas* 'shepherd' became **herdijás*, whence Gothic *hairdeis*. The twentieth-century Indologist Franklin Edgerton expanded Sievers' Law to include the liquids and nasals (leading to the term "Sievers-Edgerton" for the law), but closer examination has revealed Edgerton's extension to be true only very limitedly, and it is found only in Indic.

3.44. Sievers' Law effects were largely undone in the daughter languages; Vedic Sanskrit is the only language that shows frequent instances of it, but already Vedic is replete with examples of analogical extensions of the law to phonetic environments where it did not originally apply, and with examples of erasure of its effects. Teasing apart the original conditioning factors has been difficult, and some believe it was not a PIE rule at all, but developed independently in some of the daughter languages. According to the now generally accepted thesis of the late Austrian Indo-Europeanist Jochem Schindler, one of the conditions was that in PIE Sievers' Law only applied when the glide began the final syllable of the word; thus a word like Vedic Skt. *vaiśvānarā-* 'belonging to all men' did not become **vaiśvānarā-*. (A *v* in Sanskrit continues a *u* in PIE.) In fact, Schindler recognized that the process is essentially identical to a phenomenon known as Lindeman's Law, named

after the contemporary Indo-Europeanist Fredrik Otto Lindeman, which states that monosyllables beginning with consonant plus glide (like **k̑uō(n)* ‘dog’) had the cluster broken up in the same way as Sievers’ Law (**k̑uūō(n)*) if the word followed a word ending in a heavy syllable.

Lindeman’s Law gave rise to so-called “Lindeman’s Law variants,” two forms of the same word whose distribution originally depended on the nature of the end of the preceding word. Descendants of both **k̑uō(n)* and **k̑uūō(n)* are preserved in Vedic Sanskrit as *śvā* and *śuvā* respectively. Sometimes one or the other variant became generalized as the only form in the language: Sanskrit eventually opted for monosyllabic *śvā*, while Greek opted for disyllabic *kúōn*. Similarly, **d̑iēus* ‘sky, skygod’ shows up in Vedic both as *dyāuṣ* and as *diyāuṣ*, but in Greek the monosyllabic form was the only one to survive, in the divine name *Zeús* (where the *z* has affected the preceding consonant, forming *z*).

For Further Reading

There is no comprehensive, up-to-date account of PIE phonology written in English. The standard reference work is in German (Mayrhofer 1986), with an introduction by Warren Cowgill. A useful collection of technical articles in English, German, and French on various aspects of the laryngeal theory is Bammesberger 1988; a similar older collection, still very important, is Winter 1965. Lindeman 1987 (better than the revised 1997 edition) is also useful, although he believes in only one laryngeal for the non-Anatolian languages. The glottalic theory is most extensively presented in the first volume of Gamkrelidze and Ivanov 1984 (see Bibliography, ch. 2). A useful critical response to the glottalic theory can be found in the first part of Szemerényi 1985.

For Review

velar	resonant	“thorn cluster”	Osthoff’s Law
voiced aspirate	syllabic resonant	long diphthong	Sievers’ Law
centum	laryngeal	“schwa secundum”	Lindeman’s Law
satem	vocalized laryngeal	voicing assimilation	variants
Roman Jakobson	coloring	Stang’s Law	
Tamaz Gamkrelidze	compensatory	Szemerényi’s Law	
Vyacheslav Ivanov	lengthening	<i>boukólos</i> rule	

Exercises

- Given the charts and accompanying discussion in §§3.4–6, reconstruct the initial stop that PIE had in the form ancestral to the forms in each of the following correspondence sets. If no gloss (translation) is given, the words mean the same as the English cognate. For English, be sure to think in terms of the pronunciation of the initial sound rather than its spelling. Ignore any material in parentheses.
 - Eng. *thin*, Lat. *tenuis*, Gk. *tanaós* ‘tall’
 - Vedic Skt. *hāvate* ‘invokes’, Avestan *zavaiti* ‘invokes’
 - Eng. *tear* (from the eye), Gk. *dákru*, Welsh *deigrlyn*

- d Vedic Skt. *dhárṣati* 'is bold', Gk. *tharséō* 'am bold', Eng. *durst*, Lithuanian *drįsti* 'to dare'
- e Lat. *pūrus* 'pure', Welsh *ir* 'fresh, green'
- f Vedic Skt. *jámbhas* 'tooth', Gk. *gómphos* 'tooth', Eng. *comb*, Lithuanian *žarbas* 'sharp edge'
- g Gk. *khleúē* 'joke', Lithuanian *gláudoti* 'to joke', Eng. *glee*
- h Armenian (*str*)-*di* 'nursling', Lithuanian *dėlė* 'leech'
- i Eng. *quern* 'hand mill', Lithuanian *girmos*
- j Avestan *sraoni-* 'buttock', Welsh *clun* 'hip', Lithuanian *šlaunis* 'hip'
- k Vedic Skt. *kāmas* 'desire', Lat. *cārus* 'dear', Eng. *whore* (note: treat *wh-* as *h-*)
- l Gk. *gólēos* 'cave', Armenian *kalat* 'lair', Lithuanian *guolis* 'lair'
- m Lat. *fodiō* 'I dig', Welsh *bedd* 'grave', Lithuanian *bedū* 'I stick, dig'
- n Eng. *burn*, Lat. *for*(*ceps*) 'fire tongs', Vedic Skt. *ghṛtām* 'clarified butter'
- o Vedic Skt. *karōti* 'does', Welsh *paraf* 'I accomplish', Lithuanian *kūrti* 'to make'
- p Vedic Skt. *pánthās* 'path', Armenian *hown* 'path', Eng. *find*
- 2 For each of the correspondence sets above in 1, pick a language (from the languages used in the charts in §§3.4–6) in which a cognate is not given and predict what the initial sound of that cognate would be. For example, in 1a, if there were a Lithuanian cognate, it would begin with *t*.
- 3 The following forms contain laryngeals. Indicate (1) what the forms became after coloring (if applicable), and (2) what the forms further became after laryngeal loss (and compensatory lengthening, if applicable). The first is done as an example.
- a **ṷreh₂ǵhos* 'briar'
(1) colored to **ṷrah₂ǵhos*, which became (2) **ṷrāǵhos*
- | | |
|---|--|
| b * <i>h₃ekʷih₁</i> 'both eyes' | g * <i>h₃erō</i> 'large bird' |
| c * <i>dhoh₁ts</i> 'one that places' | h * <i>bheh₂mi</i> 'I say' |
| d * <i>h₂oǵmos</i> 'track, path' | i * <i>sneh₁si</i> 'you sew' |
| e * <i>protisth₂os</i> 'standing fast' | j * <i>neh₂us</i> 'boat' |
| f * <i>bhuuh₂tos</i> 'having come into being' | k * <i>dhedheh₁eti</i> 'he will place' |
- 4 The following forms were affected by the rule of voicing assimilation. State the outcome.
- | | | |
|--------------------------------|---------------------------------|---------------------------------|
| a * <i>h₂eǵ-tō-</i> 'driven' | c * <i>kʷid-kʷe</i> 'something' | e * <i>bhid-tō-</i> 'split' |
| b * <i>junég-si</i> 'you yoke' | d * <i>o-sd-os</i> 'branch' | f * <i>tiǵ-ter-</i> 'worshiper' |
- 5 The following forms contain resonants, laryngeals, and high vowels. Apply the PIE rules of syllabification to figure out which of the resonants and laryngeals should be syllabic or vocalized, and which of the high vowels should be non-syllabic (i.e., glides). Use appropriate diacritic marks to indicate which resonants are syllabic and which high vowels are glides.
- | | | |
|-------------------------------|-----------------------------|------------------------------|
| a * <i>h₂rnbhi</i> 'with men' | c * <i>k̑mtom</i> 'hundred' | e * <i>h₂sntieh₂</i> 'being' |
| b * <i>h₂erh₃trom</i> 'plow' | d * <i>h₂luHon-</i> 'young' | f * <i>ǵhēimno-</i> 'wintry' |

6 Observe the following forms:

- a **h₂ek^w*-*eh₂* > Lat. *aqua* 'water' (cp. **sal-* > Lat. *sal-* 'salt')
 **h₂ek^w*-*lo-* > Old Norse *ægir* 'sea' (cp. **meh₁-ri-* > Old Norse *mærr* 'famous';
 **ghan-* > Old Norse *gana* 'to gape')
- b **e-ǵneh₃-s* > Gk. *é-ǵnō-s* 'you knew' (cp. **udōr* > Gk. *hudōr* 'water')
 **ǵnēh₃-s-* > Hittite *gane-š-* 'recognize' (cp. **uēǵ-ti* > Hittite *wek-zi* 'demands';
 **uēdōr* 'waters' > Hittite *widār*)
- c **uēh₂-nti* > Hittite *wahḫanzi* 'they turn' (cp. **albh-* > Hittite *alpi-* 'cloud')
 **uēh₂-ti* > Hittite *wēḫzi* 'turns'
- d **h₃est-e-* > Gk. *osté-on* 'bone' (cp. **o-sd-os* > Gk. *ózos* 'branch')
 **h₃ēsto/o-* > Hittite *heštāš* 'bone-house' (cp. **h₃erbh-* > Hittite *ḫarapp-*
 'separate')

How do these forms show that coloring only affected short *e*?

4 Proto-Indo-European Morphology: Introduction

The Root and Indo-European Morphophonemics

4.1. Morphology is the study of the rules governing word-formation and inflection; the term also refers to the set of rules themselves in a given language. Words consist of one or more *morphemes*, the smallest meaningful units in a language. Morphemes can be whole words (e.g. Eng. *bed*, *succotash*, *sarsaparilla*) or parts of words such as affixes (e.g. the Eng. prefix *un-* and suffix *-ed*). In many languages, some morphemes can appear in different forms called *allomorphs* depending on their phonetic or morphological context (e.g. the Eng. prefix *in-* 'not' can appear as *in-*, *im-*, *il-*, *ir-*, as in *in-credible*, *im-perfect*, *il-logical*, *ir-replaceable*). The *morphophonemics* of a language is the set of rules determining the distribution of allomorphs.

A *root* is a morpheme from which semantically related words can be derived. The root itself does not usually exist as an independent form, but carries the semantic core of any word derived from it. In English, for example, the words *commit*, *emit*, *transmit*, *remission*, and *missive* are all derived from a root *mit* (borrowed from Latin) that conveys the basic meaning 'send'. When reconstructing the vocabulary of PIE, typically it is roots that are reconstructed in the first instance (see below, §4.11, for more on this). Since they did not stand alone, they are conventionally cited with an added hyphen (e.g. **sed-* 'sit'), indicating that suffixes had to be added to form free-standing words. Besides the attachment of suffixes (and sometimes, but more rarely, prefixes and infixes – affixes added into the middle of a root), word-formation in PIE often required modification of the shape of the root itself, in ways to be discussed further below.

Unlike sounds, morphemes do not necessarily change in an exceptionless and regular way over time. Changes may sometimes affect whole morphological systems, but very commonly they affect only individual words, as when the old irregular plural *kine* was replaced by the regular plural *cows* in English. Changes that affect individual words are by their nature sporadic, the results of such processes as analogy, paradigm leveling, and folk-etymology.

As with the discovery of the sounds of PIE, the determination of its morphemes and morphology was a long process, and is still going on today. The theory of ablaut (see below), in particular, took many decades to establish because of difficulties in reconstructing the PIE vowels. A guiding principle in historical morphology is that one should reconstruct morphology based on the *irregular* and *exceptional* forms, for these are most likely to be archaic and to preserve older patterns. Regular or predictable forms (like the Eng. 3rd singular present *takes* or the past tense *thawed*)

are generated using the productive morphological rules of the language, and have no claim to being old; but irregular forms (like Eng. *is*, *sang*) must be memorized, generation after generation, and have a much greater chance of harking back to an earlier stage of a language's history. (The forms *is* and *sang*, in fact, directly continue forms in PIE itself.)

The Root

Canonical shape of the root

4.2. Every language has a particular "feel," a characteristic cut to its jib, that stems from phonological properties shared by the words in the language. These properties are a product of the language's phonological inventory, of its rules regarding combinations of sounds (called *phonotactic* rules), and of the basic structure of its morphemes.

The structure of most PIE roots can be boiled down to a single template, **CeC-*, where C stands for any consonant and *e* is the fundamental vowel (on which more below). This template could be modified in certain ways, especially by adding consonants either at the beginning or the end to form consonant clusters. Most commonly, a resonant could occur on either side of the vowel, resulting in roots of the shape **CReC-*, **CeRC-*, and **CReRC-*. (Both *i* and *u* can function as resonants following the *e*; recall §3.14.)

The following are some examples of PIE roots, arranged by structure (remember that voiced aspirates such as *bh* count as single consonants):

- *CeC-* **pet-* 'fly', **ped-* 'foot', **dheg^wh-* 'burn' (remember that *dh* and *g^wh* are single consonants), **seu-* 'press out juice', **bel-* 'strength', **h₁es-* 'be', **deh₃-* 'give', **ues-* 'buy, sell', **leg^h-* 'lie down', **sem-* 'one'
- *CReC-* **dhuer-* 'door', **sneh₃-* 'sew', **tieg^w-* 'revere', **suep-* 'sleep', **smei-* 'smile', **gneh₃-* 'know', **kleu-* 'hear', **sreu-* 'flow'
- *CeRC-* **dheigh-* 'shape with the hands', **derk-* 'see', **melg-* 'wipe', **meldh-* 'speak solemnly', **gembh-* 'bite', **h₃erbh-* 'change social status', **neh₃t-* 'buttocks'
- *CReRC-* **gbrendh-* 'grind', **kreuh₂-* 'gore', **sueh₂d-* 'sweet', **mleuh₂-* 'speak'

4.3. Roots could also have any of the basic structures above preceded by *s*. Some examples include **spek-* 'see', **steg-* 'cover', **sneig^wh-* 'snow', **sreu-* 'flow', and **strenk-* 'tight'. A curious fact about such roots is that they sometimes appear without the initial *s*-, for reasons still not understood; these are called *s-mobile* roots. An example is the root **spek-* just mentioned: the *s*- is present in such forms as Av. *spasieiti* 'looks at' and Lat. *speciō* 'I look at', but absent in Ved. *páśyati* 'sees'. Similarly, from **steg-* we have on the one hand Gk. *stégo* 'I cover', but on the other Lat. *toga* 'toga' (< **covering*). The phenomenon was sporadic even within one and the same language: alongside the *s*-less Vedic form *páśyati* above, there is a causative form that has the *s*-, *spāśáyate* 'makes seen, shows', as well as a derived

noun *spás-* 'spy'. Such roots are often written with the *s* in parentheses: **(s)pek-*, **(s)teg-*.

As will be recalled from forms like **tkei-* 'settle' and **pter-* 'wing' in §3.23, and by the existence of word-initial "thorn" clusters (§3.25), a few roots began with a cluster consisting of two stops.

Roots with laryngeals

4.4. Many roots, when first reconstructed in the nineteenth century, were seen to have a shape that did not adhere to the canonical structure. These mostly fell into two categories: vowel-initial roots, such as **ant-* 'front' and **od-* 'smell', and the so-called "long-vowel" roots ending in a long vowel, like **dhē-* 'put', **pā-* 'protect', and **dō-* 'give'. Thanks to the laryngeal theory, however, all these can be shown to be of the normal type. The vowel-initial roots originally began with a laryngeal that colored the adjacent vowel and later disappeared; thus **ant-* and **od-* were once **h₂ent-* (type **CeRC-*) and **h₃ed-* (**CeC-*). The long-vowel roots once ended with a laryngeal that also disappeared, but with coloring and compensatory lengthening; thus **dhē-*, **pā-*, and **dō-* were originally **dheh₁-*, **peh₂-*, and **deh₃-*, all of them thus also of the **CeC-* type.

4.5. While this sort of laryngeal "rewriting" can be done purely for theoretical purposes (out of the wish to reconstruct an internally consistent system), in many of these cases there is direct evidence to back it up. We saw in §§3.19–20 how the laryngeals in **h₂ent-* and **peh₂-* (traditionally **ant-* and **pā-*) are preserved in Hittite as *h* in the words *hant-* 'front, forehead' and *pah₂-š-* (with a suffix) 'protect'. More difficulty is posed by roots traditionally reconstructed as beginning with *e-*, such as **es-* 'be'; this can be recast as **hes-*, but since **h₁* disappeared before a vowel everywhere (including Anatolian), evidence for the laryngeal here must come from indirect evidence. Such indirect evidence does in fact exist for many of these roots (an example is left for the exercises). Specialists differ on whether to rewrite all traditionally vowel-initial roots in this way when direct evidence of a laryngeal is lacking; the tendency is to add the laryngeal regardless, for the sake of structural uniformity.

4.6. A further class of roots, mostly beginning with a resonant in traditional reconstructions, have descendant forms with an initial vowel before the resonant in Greek and often Armenian and Phrygian (when there is evidence; not many remains of this language are preserved). Thus corresponding to Vedic Skt. *rudhirá-* 'red', *nár-* 'man', and *méh-* 'urinate' are Greek *eruthrós*, *aner-* (Armenian *ayr*, Phrygian *anar*), and *omeikh-*. In earlier scholarship, the initial vowel was simply called "prothetic" – added on – and left unexplained. It has since been shown that these vowels in fact descended from laryngeals that had vocalized in Greek, Armenian, and Phrygian. For the three roots represented by these words, we now reconstruct **h₁reudh₁-*, **h₂ner-*, and **h₃meigh_h-*.

Independent evidence for the laryngeal in such roots (when the roots are not preserved in Anatolian) comes from certain compounds. Sanskrit has an adjective *sūnára-* 'mighty, fortunate', literally 'having good heroes' or 'good in manliness'; it is a compound of the prefix *su-* 'good' and the stem *nár-* that we just saw. The *u* of the prefix is normally short, and its length in *sūnára-* can only be explained by the laryngeal beginning the word for 'man': the sequence **su-h₂ner-* became **sūner-* after laryngeal loss with compensatory lengthening.

Structure of roots with laryngeals

4.7. The bulk of roots with laryngeals fall into the three types just introduced – *CeH-, *HeC-, and *HReC. The mirror image of the last type, *CeRH, is also common, as in **uemh₁*- ‘vomit’, **terh₂*- ‘cross over, overcome’, and **kelh₂*- ‘cry out’. Less commonly, the laryngeal neighbored a stop, as in **h₃bhel-* ‘be of use’ and **pleth₂*- ‘broad’. In all these cases, the laryngeal was either the first or last consonant of the root. Some roots contained a laryngeal before the final consonant, like **dheh₃s-* ‘sacred’ and **neh₃t-* ‘buttocks’.

Roots ending in a laryngeal are often called *set* roots (pronounced like Eng. *sate*), a term from traditional Sanskrit grammar, and those without such a laryngeal are called *aniṭ*. The terms *set* and *aniṭ* mean ‘with *i*’ and ‘without *i*’, referring to the *i* that is the Sanskrit reflex of a vocalized laryngeal (cp. §4.18; see also §10.36).

Roots with *a* as fundamental vowel

4.8. A small but significant number of roots had *a* rather than *e* as the fundamental vowel. Examples include **mad-* ‘be drunk’, **nas-* ‘nose’, **sal-* ‘salt’, and **kas-* ‘gray’. For reasons that are debated, initial *k-* is particularly common in this class of roots, as in **kadh-* ‘protect’, **kamp-* ‘bend’, and **kan-* ‘sing’. As just discussed, many roots traditionally reconstructed with initial *a-*, such as **ag-* ‘drive’, are now usually thought to have begun originally with **h₂e-* (§4.4 above). For more on the behavior of these roots, see §4.17 below.

Root-structure constraints

4.9. Certain classes of consonants rarely or never co-occur within a given PIE root. There are not many securely reconstructible roots containing two plain (unaspirated) voiced stops (type **bed-*) or a voiceless stop and a voiced aspirate (type **bhet-* or **tebh-*, although the second of these is commonly found if preceded by an *s*, so **stebh-*). The source of these constraints is unknown, although similar constraints are known from other language families. See also §3.10.

Root “extensions” and “enlargements”

4.10. It is not uncommon for roots to appear with extra phonetic material (one or two sounds) added on to them, generally without any discernible change to the meaning of the root. These additional sounds are called “extensions” or “enlargements” (or “determinatives” in older literature). The root *(*s*)*teu-* ‘push, hit, thrust’, for example, appears extended or enlarged as *(*s*)*teu-k-*, *(*s*)*teu-g-*, and *(*s*)*teu-d-* (reflected respectively e.g. in Gk. *tíkōs* ‘hammer’, Eng. *stoke*, and Ved. *tudātī* ‘beats’). The source and function of these extensions are not known.

4.11. Although most of the reconstructed PIE lexicon is in the form of roots, we can also reconstruct many whole words. Most are derived from known roots, but some words, apparently belonging to a very ancient layer of IE vocabulary, cannot be (at least not uncontroversially), e.g., **seh₂uōl* ‘sun’, **dhugh₃tēr* ‘daughter’, **gheluneh₂* ‘chin’, **ag^wnos* ‘lamb’, and **uortokos* ‘quail’. A few, like the word for ‘apple’, **abel-*, and the word for ‘ax’, **pelekus*, have a shape that seems un-Indo-European

and are thought by some to be prehistoric borrowings from non-IE languages. Such proposals are not unreasonable but can rarely be evaluated. (The case of **pelekus* is instructive: it was long thought to be a borrowing from a Semitic source akin to Akkadian *pilakku(m)* or *pilaqu(m)*, but the Akkadian word was later shown to mean 'spindle', not 'ax'!)

Ablaut

4.12. Since *e* was the fundamental vowel of most PIE roots (§4.2), this is the vowel conventionally used when citing a root; so an IE dictionary entry for the root meaning 'sit' would have the headword **sed-*. Under certain conditions, however, this *e* could be replaced by other vowels. Specifically, a root could appear instead with short *o* (**sod-*), long *e* (**sēd-*), long *o* (**sōd-*), or no vowel at all (**sd-*). Broadly, the choice of vowel was determined by the type of word derived from the root. Different verb tenses, for example, called for different vowels in the root, as did different nominal (noun) inflections.

These changes in the root vowel constitute the system of vocalic alternations called **ablaut** (or **apophony**). Ablaut was central to PIE morphology, and its effects are still with us today. Consider the English forms *sing sang sung song*: between the consonants *s-* and *-ng* a different vowel appears depending on whether the form is a present-tense verb, a past-tense verb, a past participle, or a derived verbal noun. The particular vowels of these words, in fact, are descended directly from ablauting vowels of PIE.

4.13. The different ablaut variants of a root are called **grades**, and are named according to the vowel that appears. The citation form of a root (with *e*) is said to be in the *e*-grade or **full grade**, the basic ablaut grade. The other ablaut grades are the *o*-grade, **lengthened e**-grade, **lengthened o**-grade, and – when no ablauting vowel appears at all – **zero-grade**.

To illustrate, consider all the ablaut grades of the root **sed-* 'sit' again, together with some of their descendant forms:

<i>e</i> -grade (full grade) <i>*sed-</i> :	Lat. <i>sed-ēre</i> 'to sit', Gk. <i>héd-ra</i> 'seat', Eng. <i>sit</i> (<i>i</i> from earlier <i>*e</i>)
<i>o</i> -grade <i>*sod-</i> :	Eng. <i>sat</i> (<i>a</i> from earlier <i>*o</i>)
zero-grade <i>*sd-</i> :	<i>*ni-sd-o-</i> 'where [the bird] sits down = nest' > Eng. <i>nest</i>
lengthened <i>e</i> -grade <i>*sēd-</i> :	Lat. <i>sēdēs</i> 'seat', Eng. <i>seat</i>
lengthened <i>o</i> -grade <i>*sōd-</i> :	OE <i>sōt</i> > Eng. <i>soot</i> (<i>*'accumulated stuff that sits on surfaces'</i>)

Roots were not the only morphemes that ablauted in PIE; many suffixes and inflectional endings also ablauted. As an example, we may list the guises taken by the stem of the Greek word for 'father', whose second syllable is an ablauting derivational suffix. The lengthened *e*-grade is seen in the nominative singular (subject case) *patēr* 'father', the *e*-grade in the accusative singular (direct object case) *patér-a*, the zero-grade in the genitive (possessive) singular *patr-ós* 'of a father', and the ordinary and lengthened *o*-grades in a compound adjective with nominative singular *eu-pátor* 'having a good father, noble' and accusative singular *eu-pátor-a*.

4.14. While theoretically any root of the normal type could appear in any ablaut grade, in practice this was not the case. Certain roots seem to have favored certain grades, sometimes appearing in one to the exclusion of all others. The root **h₂kous-* ‘hear’ (Gk. *akoúō* ‘I hear’, Goth. *hausjan* ‘to hear’, Eng. *hear*), for example, appeared only in the o-grade, as did the adjective **bhos-o-* ‘naked’ (Eng. *bare*) and the noun **uobhs-eh₂* ‘wasp’ (Avestan *vaŋša-ka-*, Eng. *wasp*). Similarly, **bhuH-* ‘grow, be’ (Lat. *fu-tūrus* ‘about to be’, Eng. *be*) probably existed only in the zero-grade in PIE. Some roots may never have made lengthened-grade forms, while others seem to have had a certain propensity for them. Thus ablaut grades were not solely determined by the type of derivative formed from the root, but were also determined lexically.

No ablaut grade is limited to any part of speech or other morphological category, and the same morphological category may require different grades for different inflectional forms within that category (as with the nominative, accusative, and genitive of the noun for ‘father’ in the Greek example cited in §4.13 above). We therefore cannot confidently ascribe any particular meaning or semantic content to any of the ablaut grades.

Formation of the zero-grade

4.15. The zero-grade deserves some special comments because the deletion of the ablauting vowel often had effects on the syllabification of neighboring resonants in a root. In roots with a resonant directly before or after the ablauting vowel, the resonant would typically become syllabic or vocalized in the zero-grade. Thus the roots **ǵhel-* ‘yellow’, **k₁ens-* ‘proclaim solemnly’, and **kemh₂-* ‘become tired’ had zero-grades **ǵhl̥-*, **k₁ns-*, and **k₁mh₂-*; and **megh₂-* ‘big’ and **nes-* ‘we’ had zero-grades **m̥gh₂-* and **ns-*. (Most examples of syllabic resonants in PIE, in fact, occur in zero-grades.) In roots with a glide before the ablauting vowel, the glide became the corresponding high vowel (as per §3.14) in the zero-grade: thus **ies-* ‘boil’ and **suep-* ‘sleep’ had zero-grades **is-* and **sup-*. (Note that a resonant preceding a diphthong stays non-syllabic in the root’s zero-grade: thus the zero-grade of **trei-* ‘three’ and **k₁ueit-* ‘white’ were **tri-* and **k₁uit-*.) Laryngeals would become vocalized: the zero-grade of **dheh₁s-* ‘sacred’ was **dhh₁s-* with a vocalized (syllabic) **h₁* (recall §3.17).

Origin of ablaut

4.16. At the time of the breakup of PIE, ablaut was a morphologically conditioned process; but some of the grades may have originally come about through sound change. (In the histories of numerous languages, speakers have reinterpreted a phonological rule or result of a sound change as a morphological rule.) Most zero-grades are in unaccented syllables, which makes it likely that the zero-grade arose by vowel loss (syncope) in unaccented syllables. However, it should be noted that if this is true, it happened well before the stage of PIE accessible by reconstruction, since we can reconstruct some forms with accented zero-grade (such as **u₁k^hos* ‘wolf’ and **h₂ftkos* ‘bear’) or consisting only of zero-grades (e.g. **suHnus* ‘son’). Accounts of the origin of the other grades are more controversial and will not be discussed here.

Ablaut of roots with root-vowel *a*

4.17. Although the evidence is sparse, it appears that roots with *a* as fundamental vowel also ablauted. The root **sal-* 'salt' had a zero-grade **sł-*, which underlies English *silt* and German *Sülze* 'pickled meat in aspic'; the root **nas-* 'nose' has lengthened-grade derivatives such as Latin *nār-ēs* 'the nostrils' and English *nose*, both from **nas-*; and the root **laku-* 'body of water' (Lat. *lacus* 'lake', Gk. *lákkos* 'pond') had an *o*-grade form **loku-* that became Scottish Gaelic *loch* 'lake'. The view that roots in *a* ablauted is not universally accepted, but these forms are difficult to explain otherwise.

Ablaut and the laryngeal theory

4.18. Ablaut was the key to deducing the existence of the laryngeals, an insight that we owe to the nineteenth-century Swiss linguist Ferdinand de Saussure. His reasoning bears relating; it is a brilliant example of straightforward but also very bold scientific thinking.

Sanskrit has a class of verbs that contain a morpheme *-na-* infixed (i.e., inserted) into the root in the present tense, such as *yunakti* '(s)he joins' (the *-ti* is the ending for the 3rd person singular). It has another class of verbs whose presents are formed with a suffix *-nā-*, such as *punāti* 'cleanses'. These classes differ systematically in the way they inflect in different morphological categories, as exemplified below:

present tense	desiderative	infinitive
yunakti 'joins'	yokṣyati 'intends to join'	yoktum 'to join'
punāti 'cleanses'	paviṣyati 'intends to cleanse'	pavītum 'to cleanse'

If we remove what we know are infixes and suffixes from the first set, we get

yuk	yok	yok
-----	-----	-----

– that is, zero-grade of the root in the first column and full grade in the other two (the *o* in Sanskrit goes back to PIE **eu*). This pattern was known also from other types of verbs.

Now the forms for the verb 'cleanses' behave oddly by comparison. The present seems to have a suffix *-nā-* before the endings rather than an infix *-na-* plus root-final consonant plus endings. In the other forms, *pav-* can go back to a full grade (it would continue PIE **peu-*), but the *-i-* added after it did not make *pavi-* look like a root form. But Saussure noticed that, at a more abstract level, the forms for 'cleanses' behaved identically to the forms for 'join'. In *yu-na-k-ti yo-k-ṣyati yo-k-tum* there was an element *-k-* before the endings; in *pav-i-ṣyati pav-i-tum* there was also an element, *-i-*, before the endings, and in *pu-nā-ti* one could think of the length of the vowel as an extra element too. Saussure reasoned that at some level, both verbs were formed in the same way, and that what seemed to be a *nā*-suffix in *punāti* was really the same *na*-infix seen in *yunakti*, plus something. He imagined there had been some segment – call it *X* – that had been the final consonant of the root for 'cleanses' and that had lengthened the infix *-na-* to *-nā-*,

and had also become *-i-* between the consonants *v* and *ʃ*, and *v* and *t*, in the other two forms. Thus he reconstructed

*pu-na-X-ti *peuX-syati *peuX-tum

exactly parallel to

*yu-na-k-ti *yeuk-syati *yeuk-tum

In this way, it was revealed that the Sanskrit *nā*-class was exactly the same as the class formed with the infix *-na-*.

4.19. On the basis of this and many other pieces of evidence, Saussure deduced the existence of the laryngeals. He claimed there were two of them, and proposed that they patterned like resonants in having syllabic reflexes when between consonants (or sounds that could be consonantal, like the *u* in **peuX-syati*). Hermann Möller, one of the few people to pick up his theory at the time and develop it further, expanded the number to three in 1879; he was joined by a third scholar, Albert Cuny, at the close of the nineteenth century. Möller's and Cuny's ideas are quite close to those standardly accepted today.

But it was not until 1927, a good half-century after Saussure's proposals, that real vindication of the whole theory came. In that year, the young Polish linguist Jerzy Kuryłowicz published his discovery that the sound *h* in the newly deciphered language Hittite appeared in many of the places that Saussure had predicted these mystery segments should have existed in PIE.

Saussure was at the ripe old age of 19 or 20 when he developed these ideas; he published his findings at age 21. What was new in his method was the application of the technique of internal reconstruction to *reconstructed PIE itself*, at the time a very radical thing to do – especially if it resulted in the positing of a set of consonants that had not survived as consonants into recorded history (as far as was then known). His starting point was the supposition that superficially different formations belonging to the same morphological categories had once been formed identically. Unfortunately, he did not live to see Kuryłowicz's work.

Morphological Categories of PIE

Word structure

4.20. The process of forming a word from a root or another word is called *derivation*, while the process of creating different grammatical forms of a given word is *inflection*. In English, prefixes like *anti-*, *con-*, *fore-*, and *hyper-*, and suffixes like *-ance*, *-ity*, *-ly*, and *-ness* are derivational morphemes; they are used to create new words. By contrast, suffixes like the plural *-(e)s* or the comparative suffix *-er* for adjectives are *inflectional morphemes*; adding them does not produce a new word, but just a different grammatical form of an already existing word.

Typically, a word in PIE consisted of three morphemes, root plus suffix plus ending, symbolized as *R + S + E*. (Prefixes and compounds will be treated below.) The suffix was a derivational morpheme, such as one used to form a present-tense

stem or an abstract noun from a verbal root. (We know of one derivational *infix* that was inserted into the root, which we saw a preview of above in §4.18.) The ending (also called the *desinence*) was the inflection, which marked the grammatical function of the word (such as the nominative singular of a noun, or the 3rd person plural of a verb). All three of these elements could and did ablaut. To take an example, the word **m̥t̥éi-s* 'of thought' consisted of the root **men-* 'think' in the zero-grade, followed by an ablauting suffix **-t(e)i-* (here in the full grade) that was used to form abstract nouns (§6.42), in turn followed by the inflectional ending **-s* of the genitive case (meaning 'of') in the zero-grade.

The suffix did not always appear overtly, in which case one speaks of *zero-suffixes*. For example, the form **dem-s* meant 'of the house' and consisted of the root **dem-* 'to build' in the *e*-grade plus the inflectional ending of the genitive singular, without an overt derivational suffix. Words could in fact be derived from one another simply by changing the ablaut or the position of the stress (or both) – a process known as *internal derivation*, much like *song* is derived from the verb *sing*. Zero-suffixes and internal derivation will be discussed in more detail in chapter 6.

Words could also be formed by *composition*, which includes prefixation and compounding. Not many PIE prefixes have been reconstructed, but one prefixal process known as *reduplication* was very common. In reduplication, a prefix was constructed consisting of a copy of the first consonant or consonant cluster of a root plus a vowel. Thus **de-deh₃-* and **ǵi-ǵnh₁-* are reduplicated forms of the roots **deh₃-* 'give' and **ǵenb₁-* 'beget', respectively. Compounding was most frequently seen in nouns and adjectives, but also occasionally in verbs; it will be discussed in the noun and verb chapters.

4.21. PIE was a richly inflected language. It possessed full sets of singular, dual, and plural endings in all three persons of the verb in several tenses, and at least eight (and possibly nine) cases in the noun, with somewhat less differentiation of cases in the dual and plural than in the singular. Adjectives and pronouns were fully inflected like nouns, as were the first four cardinal numerals. The functions of the case-endings of the noun, the personal endings of the verb, and the other verbal grammatical categories such as tense, voice, and mood, will be explained in detail in the two chapters to follow.

Thematic and athematic inflection

4.22. PIE nouns, adjectives, and verbs can be divided into two basic groups based on their inflectional patterns: those that had an ablauting short vowel, *e* or *o* (in shorthand, *-e/o-*), directly before the inflectional endings (the case-endings in nouns or adjectives, the personal endings in verbs), and those with no such vowel. This ablauting vowel is called the *thematic vowel*; words that inflect with it are termed *thematic*, and those inflecting without it are *athematic*.

For example, the Greek word *klṓps* 'thief' consists of the Greek (and IE) root *klep-* (in the lengthened *o*-grade) followed directly by the inflectional ending *-s*, which marks the nominative singular (or subject case). *Klṓps* is therefore an athematic nominative singular. By contrast, the Greek noun *nómos* 'law, custom' has the vowel *-o-* between the root (*nom-*, *o*-grade of *nem-*) and the ending. *Nómos* is therefore a thematic nominative singular.

4.23. Athematic declensions and conjugations are on the whole more complex than their thematic counterparts. Within athematic paradigms, there are alternations in ablaut and changes in the position of the accent; these will be fully discussed in the next two chapters. Athematic inflection appears to belong to a more ancient layer of IE nominal and verbal derivation than thematic inflection, and many athematic formations are moribund by the time they are attested, even in languages as archaic as Vedic Sanskrit. In the observable course of the histories of the older IE languages, thematic paradigms as a rule are outstripping and replacing athematic ones, giving thematic formations the appearance of young upstarts taking over the field. For example, the athematic Greek verb *seû-tai* 'is hunted, chased', found only in Homer and some other early Greek poetry, was remade as – and replaced by – a thematic verb *seû-e-tai*. Similarly, prehistoric Greek had an athematic noun with stem **thes-* meaning 'god'; as a free-standing noun it was replaced by the thematic stem **thes-o-*, eventually becoming (after sound changes) Classical Greek *theós*. The older athematic stem survived only as a frozen member of compounds like *thés-phatos* 'decreed by god'. Thematic paradigms, unlike athematic ones, did not exhibit internal shifts in accent or alternations in ablaut grade and were thus arguably simpler for speakers to handle, which may partly explain their spread. The full reasons, however, are surely more complicated, as "simplicity" is a subjective and notoriously misleading yardstick in historical linguistics.

4.24. We are lucky to have a wealth of good comparative material on which to base our reconstructions of PIE morphology. The precise agreements among the ancient languages in nominal and verbal derivation and inflection are at times astounding, and leave little room for doubt about a great many aspects of the PIE morphological system. Many other aspects, however, are still the subject of controversy, and claims about them are more tentative. The next three chapters will provide an overview of the standard reconstruction of the PIE verbal, nominal, and pronominal systems, and the comparative evidence upon which it is based. Usually, only a selection of comparative data will be adduced in support of a given reconstruction so as not to swell the sizes of these chapters unmanageably; but where that would have sacrificed clarity or blunted the understanding of an important issue, it was thought better to give more information rather than less.

The material in the next few chapters is at any rate rather extensive, and is not meant to be learned all at once. Some of it is included primarily for reference. If it appears bewildering, the exercises can be used for orientation, since they focus on the most important points to be digested. As chapters 9 and following are worked through, with their sections on historical morphology, additional material from chapters 5–7 can be drawn in.

For Further Reading

Most studies in the realm of PIE morphology are devoted to specific aspects of it, such as noun or verb formation; references to such studies are given in the following chapters. No good general works on root structure or ablaut are found in English. For those who read French, influential is Chapter IX of Benveniste 1935 on root structure and root "enlargements." Kurylowicz 1956, a highly complex work, will be impenetrable to introductory students, but is an important and famous treatment of ablaut. Ferdinand de Saussure's pio-

neering work on laryngeals was published as Saussure 1879, and Kurylowicz's vindication can be read in Kurylowicz 1927. Note also Anttila 1969, which concerns roots that show both *CRēC* and *CēRC* forms and various related ablaut issues.

For Review

Know the meaning or significance of the following:

root	ablaut	Albert Cuny	composition
s-mobile	grade	Jerzy Kurylowicz	thematic vowel
set	Ferdinand de Saussure	derivation	thematic inflection
anŷ	Hermann Möller	inflection	athematic inflection

Exercises

- 1 Recalling §3.23, indicate which of the following would be well-formed PIE roots:

a **streibh-* b **rdeu-* c **bhal-* d **h₂seup-* e **nregh-*

- 2 Which of the following would violate the root-structure constraints in §4.9?

a **kebh-* b **skebh-* c **kep-* d **gep-* e **beg-*

- 3 The following are some reconstructed roots in the full grade. Provide the zero-grade and the lengthened *o*-grade.

a * <i>ǵheu-</i> 'pour'	e * <i>bherǵh-</i> 'bright'	l * <i>pleh₁-</i> 'fill'
b * <i>h₂ekʷ-</i> 'see'	f * <i>keh₂d-</i> 'care'	j * <i>ǵembh-</i> 'bite'
c * <i>k₁uēn-</i> 'holy'	g * <i>jes-</i> 'boil'	k * <i>sengʷh-</i> 'sing'
d * <i>skel-</i> 'dry out'	h * <i>dheb-</i> 'thick'	l * <i>h₁reudh-</i> 'red'

- 4 For each of the roots in 3, indicate what basic type they belong to (**CēC*, **CRēC*, **CēRC*, etc.). Ignore initial *s-* in consonant clusters.

- 5 Identify the ablaut grade that the following root forms appear in:

a * <i>sōd-</i> 'sit'	c * <i>uoid-</i> 'see'	e * <i>ǵhel-</i> 'yellow'
b * <i>m₁-</i> 'waste away'	d * <i>stēu-</i> 'praise'	f * <i>diḱ-</i> 'point'

- 6 The Latin cognate of English *foam* is *spūma*. Explain the *f*-/sp- correspondence using material from this chapter as well as §3.4.

- 7 The following are older, "pre-laryngealistic" reconstructions of roots that are nowadays considered to have contained a laryngeal and the root vowel *e*. Rewrite the reconstructions using laryngeals.

a * <i>kā-</i> 'love'	c * <i>snē-</i> 'spin, sew'	e * <i>okʷ-</i> 'see'
b * <i>ank-</i> 'bend'	d * <i>uēr-</i> 'water'	f * <i>bhlāg-</i> 'strike'

8 It was noted in §4.5 that the root for 'be' in PIE was **h₁es-*.

- a Give its zero-grade.
- b Vedic Sanskrit has a prefix *a-* 'not, un-' and the present participle of the verb 'to be' is *sat-* 'being'. Given that the *a* in both these forms comes from PIE **h₁*, and the present participle was formed from the zero-grade of the root, what would the PIE ancestors of *a-* and *sat-* have been?
- c Vedic Skt. has a word *āsat-* 'monster', a compound of *a-* and *sat-* above. Provide a historical explanation for why the first *a* is long rather than short.

9 For each of the following English words, identify the morphemes, and specify which are derivational and which are inflectional.

- | | | |
|---------------|---------------|--------------------|
| a capitalizes | d angriest | g disproportionate |
| b happier | e superpowers | h recklessness |
| c carrier | f unbending | i insurmountably |

10 The following are reconstructed words in PIE. Each consists of several morphemes, separated from each other by hyphens. One of the morphemes is a form of a root. Using your knowledge of ablaut and root structure, identify the root that the word is derived from, in the *e*-grade. Assume that long vowels are not induced by laryngeals.

Example: In **bhe-bhoid-e* '(s)he has split', the underlying root must be **bheld-* (*bhe-* and *-e* cannot be forms of roots because of their structure).

- | | |
|---|--|
| a <i>*h₁-dhg^hhi-to-m</i> 'imperishable' | e <i>*g^hh₁-t-osjo</i> 'of the slain one' |
| b <i>*ley-o-s</i> 'grain' | f <i>*yōt-eno-s</i> 'enraged' |
| c <i>*dor-u</i> 'wood' | g <i>*de-dork-e</i> '(s)he has seen' |
| d <i>*yid-mē</i> 'we know' | |

11 Identify the following forms as thematic or athematic. Each morpheme is separated by a hyphen.

- | | |
|--|--|
| a <i>*bher-e-si</i> 'you carry' | d <i>*yīh₁-o-s</i> 'man, hero' |
| b <i>*kun-ēs</i> 'of the dog' | e <i>*e₂-steh₂-t</i> '(s)he stood' |
| c <i>*m₁-tēi-s</i> 'of thought' | f <i>*k^wel-o₁-nti</i> 'they turn' |

5 The Verb

The Structure of the PIE Verb

5.1. Verbs in PIE and in most of the ancient IE daughter languages were inflected in dozens of forms. In this they were vastly different from Modern English verbs, barely any of which has more than five forms (e.g. *sing sings singing sang sung*). As stated in the preceding chapter, the grammatical categories that were distinguished in PIE verb inflection are **person** (first, second, and third), **number** (singular, dual, and plural), **tense**, **voice**, and **mood**. (On verbal aspect, see §5.10 below.)

Tense

5.2. Verbs in PIE could inflect in the **present tense** ('I go, am going'), the **imperfect tense** ('I was going', expressing ongoing or background action in past time), and the **arist tense** ('I went', expressing one-time or completed action in past time). Traditional scholarship also recognizes a fourth tense, the **perfect**, but this is now viewed as a stative (see §5.49 below) that secondarily acquired use as a **resultative past tense**: 'I am in a state of having gone, I have gone'. There was probably also a **pluperfect**, a past of the perfect, though this is debated. Some further believe that PIE had a **future** (see §§5.35ff. below). Each of these tenses could distinguish various voices and moods (more limitedly, it appears, in the perfect), which are discussed in §§5.50ff.

5.3. The actual forms of the tenses and moods were made from **tense-stems**, of which there were three: the **present stem**, **arist stem**, and **perfect stem**. There do not appear to have been conjugational classes in PIE as there were in, say, Latin. Rather, in one group of verbs, called **primary verbs**, the tense-stems were formed directly from the root. In another group, called **derived verbs**, the tense-stems were created secondarily by means of productive suffixes to express particular types of action or shades of meaning. These verbs included causatives, iteratives, desideratives, and denominatives, all of which will be discussed in due course.

Not every verb could form all three tense-stems. Quite a few did not form perfects, for example, and derived verbs only had present stems in PIE. (The daughter languages have often independently innovated additional tense-stems for these verbs.) A few verbs exhibited **suppletion**, meaning they had different tense-stems formed from different roots. Examples of suppletive paradigms in English are *go/went* and *be/is/were*.

Voice

5.4. The voice (also called *diathesis*) of a verb indicates the role that the subject plays in the action. Two voices were distinguished in verbal inflection, active and middle. (On the passive, see below.) The difference in meaning between these two voices in PIE is not fully clear. In traditional grammatical usage, active means that the subject is doing the action rather than being acted upon, while middle means the subject is either acting upon itself or is in some other way “internal” to the action.

This rough guideline works reasonably well for verbs that could inflect in either voice. We can illustrate this with some active/middle verb pairs in Hittite. Sometimes the middle meant the same as the active but differed in not taking a direct object; in other words, the middle could express the intransitive of a transitive active, e.g. transitive active *irhāizzi* ‘sets an end to’, intransitive middle *irhāitta* ‘comes to an end’. Second, the middle could express the reflexive sense of the active, where the subject acts on itself, e.g. active *nāi* ‘leads’, middle *neyari* ‘leads oneself, turns (oneself) to’. Finally, the middle could have a reciprocal sense, e.g. active *zahhīyawem* ‘we fight (someone)’, middle *zahhīyawaštati* ‘we fight each other’. These uses of the middle are comparable to the reflexive in modern Romance, Slavic, and German, e.g. German *sich wenden* ‘turn (oneself) to’, *sich streiten* ‘argue (with one another)’.

5.5. But in many other cases, the distinction between active and middle inflection was purely a formal one: there were some verbs that inflected only in the active and others only in the middle, without clear difference in meaning. Verbs having only middle inflection are often called **middle verbs**. (Students familiar with Latin can think of these as equivalent to the Latin deponent verbs – active in meaning but having only passive endings, which come historically from the PIE middle.) It is not fully clear whether their middle inflection stemmed from some aspect of their meaning, or whether it was purely arbitrary. On the one hand, as a group these verbs do tend to express various “internal” or intransitive notions like spatial movement, position of rest, emotions, sensory perception, speaking, giving off sound or light, and changes of state. (In technical terms, these are mostly “unaccusative” verbs.) However, active verbs could also express such notions. Compare the representative list of middle verbs in the left-hand column below with the active verbs on the right; the meanings are unaccusative in both cases, sometimes identical:

Middle

Hittite *iyattari* ‘goes’
 Vedic Sanskrit *śāye* ‘lies’
 Latin *ueretur* ‘fears’
 Greek *dérketai* ‘sees’
 Greek *eúkhetai* ‘proclaims’
 Latin *moritur* ‘dies’

Active

Greek *eīsi* ‘goes’
 Vedic Sanskrit *sīdati* ‘sits’
 Greek *khaírei* ‘rejoices’
 Latin *audī* ‘hears’
 Hittite *memai* ‘says’
 Hittite *ḫarakzi* ‘dies’

5.6. Further complicating our understanding of the middle is that some verbal roots had active inflection in one tense or mood, but middle in another. For instance, the Greek verb ‘learn’ has an active present *manthānō* ‘I learn’ but a middle future *mathēsomai* ‘I will learn’; this is a common pattern. The Old Irish verb *ad-ciú* ‘I see’ was active in the present indicative but middle in the present subjunctive (*ad-cear*);

and the Tocharian A active present indicative *yāṣ* ‘goes’ has a middle present participle *ymān* ‘going’. (Again for students who know Latin, the Latin “semi-deponent” verbs, such as present [active] *audēō* ‘I dare’, perfect [passive] *ausus sum* ‘I have dared’, are comparable.)

5.7. The middle could also express the passive voice, which indicates that the subject is acted upon by someone else: ‘is being fought’, ‘was washed’. A tradition of scholarship rejects positing a passive voice for PIE because there was no separate set of passive endings. But all the daughter languages that have a separate passive conjugation have developed it in whole or in part from the PIE middle endings, and it seems best to regard the middle as having been, in fact, a *mediopassive* or *middle-passive* – capable of expressing either voice depending on the context.

Mood

5.8. The mood in which a verb appears expresses the speaker’s attitude or stance taken towards the action – whether (s)he is asserting that it is factual, or indicating a wish that it were or were not true, or reporting the action second-hand, or indicating a contrafactual condition. PIE possessed four moods: indicative, imperative, subjunctive, and optative.

5.9. All the daughter languages agree that the indicative was used to express matters of fact, and the imperative was used for issuing commands. The function of the subjunctive in PIE is less clear, but it was probably at least in part a future tense; see §5.52 for a full discussion. Finally, the optative was used to express wishes and various other non-factual modalities.

Aspect

5.10. Aspect is a grammatical category that refers to the type of action indicated by a verb. Actions can be done once or repeatedly, to completion or not, or be ongoing with neither a true beginning nor end. It is on the whole unclear what the aspectual system of PIE was; in part the difficulty lies in figuring out the niceties of aspectual differences in the earliest texts in the daughter languages. For example, part of the opening of the famous Rig Vedic hymn 1.32 reads: “I (will) tell now the manly deeds of Indra, the foremost of which he did bearing a cudgel. He slew the serpent, drilled through to the waters, (and) split the belly of the mountains.” The underlined verbs in the Sanskrit original are in the perfect, imperfect, perfect, and imperfect, respectively; yet any difference in aspectual sense eludes us.

According to the generally accepted view, the imperfect and aorist were distinct aspectually, the imperfect expressing incomplete or ongoing action in past time (imperfective aspect), the aorist indicating completed or punctual (one-time) action in past time (perfective aspect).

5.11. The persons of the verb were expressed with suffixes called **personal endings**. A complete set consisted of nine forms, the three persons in each of three numbers (singular, dual, and plural). PIE had several sets of personal endings. The

most fundamental distinction was among those of the active voice, middle voice, and the perfect. Additionally, for each of the two voices active and middle, non-past or **primary** endings and past or **secondary** endings were distinguished. The primary endings were used for the present tense and the subjunctive mood, while the secondary endings were used for the two past tenses (imperfect and aorist) and for the optative mood. In sum there were five sets of personal endings with the following distribution:

<i>Personal endings</i>	<i>Where used</i>
primary active	present indicative active, active subjunctives
secondary active	imperfect and aorist indicative active, active optatives
primary middle	present indicative middle, middle subjunctives
secondary middle	imperfect and aorist indicative middle, middle optatives
perfect	perfect

As an illustration, consider the following 3rd singular forms of the PIE verb 'turn': present indicative active **k^wele-ti* (primary active ending), imperfect indicative active **k^wele-t* (secondary active ending), present indicative middle **k^wele-tor* (primary middle ending), imperfect indicative middle **k^wele-to* (secondary middle ending), perfect **k^wek^wol-e* (perfect ending).

Not all of the endings can be reconstructed with equal certainty; the most secure are those of the singular and the 3rd person plural. We shall treat the active and middle personal endings first; the perfect endings will be discussed in §5.47.

Active endings

5.12. Below is a selection of the comparative evidence used to reconstruct the PIE primary and secondary active endings, whose reconstructions are given in the right-hand column. In these and the following tables, only forms relevant for the reconstructions are given, and the information is by no means exhaustive.

PRIMARY (NON-PAST) ACTIVE ENDINGS

	Vedic Sanskrit	Greek	Gothic	PIE
sg. 1	<i>é-mi</i> 'I go'	<i>ei-mí</i> 'I am' etc.	<i>baíra-m</i> 'I bear' etc.	<i>*-mi</i> , <i>*-h₂¹</i>
2	<i>é-ši</i> 'you go'	<i>es-sí</i>	<i>baíra-s</i>	<i>*-s₁</i>
3	<i>é-ti</i> '(s)he/it goes'	<i>es-tí</i>	<i>baíra-þ</i>	<i>*-ti</i>
du. 1	<i>i-vás</i> 'we both go'		<i>baír-os</i>	<i>*-ye-</i>
2	<i>i-thás</i> 'you both go'	<i>es-tón</i>	<i>baíra-ts</i>	<i>*-to-</i>
3	<i>i-tás</i> 'they both go'	<i>es-tón</i>		<i>*-to-</i>
pl. 1	<i>i-más(i)</i> 'we go'	<i>ei-més</i> (Doric)	<i>baíra-m</i>	<i>*-me-</i>
2	<i>i-thá</i> 'you (pl.) go'	<i>es-té</i>	<i>baíra-þ</i>	<i>*-te(-)</i>
3	<i>y-ánti</i> 'they go'	<i>enti</i> (Doric)	<i>baíra-nd</i>	<i>*-(é)nti</i>

¹ Not illustrated here; see below, §5.13.

SECONDARY (PAST) ACTIVE ENDINGS

Note: the Vedic and Greek forms are imperfects meaning 'was/were carrying'; the Gothic forms are present optatives meaning 'would carry'. The vowels *a-* and *e-* at the beginning of the Sanskrit and Greek forms are explained in §5.40.

	Vedic Sanskrit	Greek	Gothic	PIE
sg. 1	<i>ābhara-m</i>	<i>éphero-n</i>		*-m
2	<i>ābhara-s</i>	<i>éphere-s</i>	<i>baírai-s</i>	*-s
3	<i>ābhara-t</i>	<i>éphere</i>	<i>baírai</i>	*-t
du. 1	<i>ābharā-va</i>		<i>baírai-wa</i>	*- ue (-)
2	<i>ābhara-tam</i>	<i>éphére-ton</i>	<i>baírai-ts</i>	*-to-
3	<i>ābhara-tām</i>	<i>éphére-ton</i>		*-teh ₂ -
pl. 1	<i>ābharā-ma</i>	<i>éphéro-men</i>	<i>baírai-ma</i>	*-me(-)
2	<i>ābhara-ta</i>	<i>éphére-te</i>	<i>baírai-þ</i>	*-te(-)
3	<i>ābhara-n</i>	<i>éphero-n</i>	<i>baírai-na</i>	*-(é)nt

NOTES ON THE ACTIVE ENDINGS

5.13. As can be seen, the primary and secondary active endings differ partly with respect to whether an *-i was added in PIE. This *-i, which marks primary active endings, has been termed the "hic et nunc" particle (Latin for "here and now"); it may be the same as the -i found attached to a number of pronominal and adverbial forms in various daughter languages, as in Greek *nūn-i* 'now'.

The reconstructions of the singular and the 3rd plural endings are uncontroversial. As for the other endings, the languages do not agree so precisely. For example, a disagreement between forms ending in -s and ending in a nasal bedevils the exact reconstruction of the 2nd and 3rd dual and the 1st plural primary endings.

The 1st singular primary ending *-mi was originally proper only to athematic presents (§§5.17ff.), while *-h₂ is the ending found in thematic presents (see §§5.25ff.) and in subjunctives (§5.51). Similarly, the 3rd plural endings *-(é)nti and *-(é)nt have the accented é only in athematic verbs.

Middle endings

5.14. The middle endings are more difficult to reconstruct than the active, partly because many of the original endings were remade in the daughter languages on the model of the active endings. In the PIE column in the charts below, only reconstructions for the singular and plural are given. The first chart contains present middles: Vedic Skt. *bharé* and Gk. *phéromai* 'I carry (for myself), I gain', Tocharian A *trānkmar* 'I say' (with a 3rd dual from another verb in Tocharian B, *nes-tem* 'both are'). The second chart contains the imperfect middles Vedic Skt. *ābhari* and Gk. *éphéromen* 'I was carrying (for myself)', and the Tocharian B preterite middle *kautamai* 'I was split':

PRIMARY (NON-PAST) MIDDLE ENDINGS

	Ved.	Gk.	Toch. A	PIE
sg. 1	<i>bhār-e</i>	<i>phéro-mai</i>	<i>trānk-mar</i>	*-h ₂ er
2	<i>bhāra-se</i>	<i>phére-ai</i> ¹	<i>trānk-tār</i>	*-th ₂ er
3	<i>bhāra-te</i>	<i>phére-tai</i>	<i>trānk-tār</i>	*-or, *-tor
du. 1	<i>bhārā-vahe</i>	<i>phéro-methon</i>		
2	<i>bhāre-the</i>	<i>phére-sthon</i>		
3	<i>bhāre-te</i>	<i>phére-sthon</i>	[nes-tem]	
pl. 1	<i>bhārā-mahe</i>	<i>phéro-metha</i>	<i>trānk-āmtār</i>	*-medhh ₂ ?
2	<i>bhāra-dhve</i>	<i>phére-sthe</i>	<i>trānk-cār</i>	*-dh(u)ye- ?
3	<i>bhāra-nte</i>	<i>phéro-ntai</i>	<i>trānk-āntār</i>	*-ro(r) ?, *-ntor

SECONDARY (PAST) MIDDLE ENDINGS

	Ved.	Gk.	Toch. B	PIE
sg. 1	<i>ābhar-i</i>	<i>epheró-mēn</i> ²	<i>kautā-mai</i>	*-h ₂ e
2	<i>ābhara-thās</i>	<i>ephér-ou</i>	<i>kautā-tai</i>	*-th ₂ e
3	<i>ābhara-ta</i>	<i>ephére-to</i>	<i>kautā-te</i>	*-o, *-to
du. 1	<i>ābharā-vahi</i>			
2	<i>ābhare-thām</i>	<i>ephére-sthon</i>		
3	<i>ābhare-tām</i>	<i>epheré-sthēn</i> ²		
pl. 1	<i>ābharā-mahi</i>	<i>epheró-metha</i>	<i>kautā-mt(t)e</i>	*-medhh ₂ ?
2	<i>ābhara-dhvam</i>	<i>ephére-sthe</i>	<i>kautā-t</i>	*-dh(u)ye- ?
3	<i>ābhara-nta</i>	<i>ephéro-nto</i>	<i>kautā-nte</i>	*-ro, *-nto

¹ Homeric (not actually attested for this verb). ² Doric *-mān*, *-sthān*, which preserve the original vowel quality *ā*.

RECONSTRUCTION OF THE MIDDLE

5.15. The most important general feature is the presence of an *-r* in the primary middle, absent in the secondary middle. This *-r* is now generally thought to have been the primary middle marker, corresponding to the **-i* of the active. Middles in Anatolian, Italic, Celtic, Tocharian, and Phrygian preserve this *-r*, but it has been replaced by the *-i* of the active in Indo-Iranian, Greek, Germanic, and Albanian (whence, e.g., the *-ai* middle endings in the Greek paradigms above). Thus in the 3rd person singular, what was originally **-to + *-r* (yielding ultimately Hitt. *-tari*, Lat. *-tur*, Old Irish *-ther*, Tocharian *-tār*, Phrygian *-tor*) was elsewhere remade as **-to + *-i* (yielding Skt. and Avestan *-te*, Gk. *-tai* [with *-toi* in some dialects], Gothic *-da*).

The element **-h₂* in the 1st person singular is surely the same as that seen in active **-h₂* in the thematic endings and the perfect (§§5.25, 5.47). The 2nd singular secondary ending was remade to **-so-* in many branches, using the *-s* from the active ending **-s(i)*, as in Gk. *ephérou* above from **e-phere-so*.

The 3rd person forms are of special interest. Regarding the 3rd singular, archaic athematic forms in Anatolian, Indo-Iranian, Celtic, and Tocharian have no *-t* (e.g. Cuneiform Luvian *zīy-ari* and Vedic Skt. *śáy-e* 'lies', both from **kei-o(-)*), suggesting that the *-t* was a later intrusion from the active. The same is true of the 3rd plural

*-nto-, for vestigial archaic middle forms point to *-ro as the earlier ending (directly attested apparently only in Tocharian B *sta-re* 'they are' from **sth₂-ro*, but compare also Vedic Skt. *śé-re* and Avestan *sōi-re* 'they lie' < **kei-ro-i* with -i added later).

The Present Stem

5.16. The present stem was used in PIE to form one primary tense, the present, and one secondary tense, the imperfect. Both of these could inflect in the active and middle voices. Along with the present tense in the indicative mood (used to express ordinary statements of fact), a present subjunctive, optative, imperative, and participle could also be formed from this stem.

Athematic presents

5.17. Like other verbal and nominal stems, present stems were either athematic or thematic (see §4.22), and there were a number of different types of each. We shall turn our attention first to the athematic presents, which were probably the older type.

5.18. Besides lacking a thematic vowel before the personal endings, the paradigms of athematic presents were characterized by changes in ablaut and typically by shifts in the position of the accent. The basic pattern is the same for all athematic presents: in the singular active, the root (or infix in the case of the nasal-infixed presents discussed below) receives the stress and is in the full grade; in the dual, plural, and in all middle forms the stress migrates rightward to the personal endings, and the root or infix is reduced to the zero-grade. Slightly different are the "Narten" presents discussed below in §5.20, but the basic distinction of a "strong" ablaut grade in the singular versus a "weaker" one in the dual, plural, and middle still obtains.

ROOT ATHEMATIC PRESENTS

5.19. The simplest and most common athematic present was the root athematic present (or root present for short), formed by adding the personal endings directly to the root. There are two types of root presents. The more common type had accented root in the *e*-grade in the singular active, and unaccented root in the zero-grade in its dual, plural, and middle. A classic example is the verb 'be', **h₁es*-; paradigms from some of the daughter languages in the singular and plural are given below, together with the PIE reconstruction on the right. (Note that Eng. *am*, *is* is a direct continuation of PIE **h₁és-mi*, **h₁és-ti*.)

	Hitt.	Ved.	Gk.	Lat.	Goth.	OCS	PIE
sg. 1	<i>ēšmi</i>	<i>ásmi</i>	<i>eimí</i>	<i>sum</i> ³	<i>im</i>	<i>jesmǐ</i>	* <i>h₁és-mi</i>
2	<i>ēši</i>	<i>ási</i>	<i>eĩ, essí</i> ¹	<i>es</i>	<i>is</i>	<i>jesi</i>	* <i>h₁és-si</i> ⁴
3	<i>ēšzi</i>	<i>ásti</i>	<i>estí</i>	<i>est</i>	<i>ist</i>	<i>jestǔ</i>	* <i>h₁és-ti</i>
pl. 1		<i>smás</i>	<i>esmén</i>	<i>sumus</i>	<i>sijum</i>	<i>jesmǔ</i>	* <i>h₁s-mé(-)</i>
2		<i>sthá</i>	<i>esté</i>	<i>estis</i>	<i>sijup</i>	<i>jeste</i>	* <i>h₁s-té(-)</i>
3	<i>ašarzi</i>	<i>sánti</i>	<i>eisí, entí</i> ²	<i>sunt</i>	<i>sind</i>	<i>sotǔ</i>	* <i>h₁s-énti</i>

¹ *eĩ* is Attic, *essí* is Homeric. ² *eisí* is Attic, *entí* is Doric. ³ Archaic Lat. *esom*. ⁴ Became **h₁esi* by the rule in §3.37.

5.20. The second type of root present, only identified in the late 1960s by the Indo-Iranianist Johanna Narten and informally termed a Narten present in her honor, has the accent on the root throughout, but an alternation between lengthened \bar{e} -grade in the singular and ordinary e -grade elsewhere. Thus from $*stew-$ 'praise' was formed 3rd sing. $*st\bar{e}w-ti$, 3rd pl. $*st\bar{e}w-nti$. Note that the 3rd plural ending is $*-nti$ instead of accented $*-enti$ in the ordinary type. The Narten presents are also called *acrostatic presents* (the term *acrostatic*, which will be properly introduced in the next chapter, means the accent remains on the root throughout the paradigm).

Ablaut of the Narten type may not have been limited to presents; see §5.43.

OTHER ATHEMATIC PRESENTS

5.21. Two other main kinds of athematic presents are known, the nasal-infix presents and reduplicated athematic presents. As these are more complex formally than root athematics, they are often called *characterized presents*. The difference was significant and had consequences elsewhere in the verbal system; see §5.42.

5.22. Nasal-infix presents. An infix is a morpheme placed inside another morpheme. One PIE infix is known, used by certain roots to form present stems. It had the shape $*-ne-$ in the full grade and $*-n-$ in the zero-grade, whence the name *nasal present* for stems containing this infix. The distribution of the ablaut grades was the same as in root presents: full grade in the singular active, zero-grade elsewhere.

The infix was inserted into the zero-grade of the root, between its last two sounds (typically a resonant or high vowel followed by a consonant): thus from $*j\bar{e}ug-$ 'yoke', zero-grade $*jug-$, the 3rd sing. nasal present $*ju-n\acute{e}-g-ti$ 'yokes' was formed (> Vedic Skt. *yunákti*), with 3rd pl. $*ju-n-g-enti$ (> Vedic Skt. *yunjānti*). Nasal-infix presents are typically active transitives, but beyond that we do not know what the meaning of the infix may once have been. Outside of Anatolian and Indo-Iranian, the infix no longer ablauts, as in Lat. *iungit* 'joins', pl. *iungunt*.

5.23. The daughter languages also have various other types of nasal presents that are ultimately related to the nasal-infix type. Two in particular may be mentioned: the suffixed $*-neu/-nu-$ presents (e.g. Vedic Skt. *tanóti* 'stretches' ~ *tanvánti* 'they stretch'; Hittite *ar-nu-mi* 'I cause to go'; Gk. *ór-nū-mi* 'I rouse, incite'); and the so-called ninth class of Indo-Iranian verbs, with an ablauting suffix $-nā/-nī-$ that goes back to $*-ne-H/-n-H-$ (that is, the nasal infix inserted before a root-final laryngeal), as in Vedic Skt. *punáti* 'cleanses' ($*pu-né-h_2-ti$) ~ *punímás* 'we cleanse' ($*pu-n-h_2-mé-$). On this last type, recall the discussion in §4.18; the Sanskrit *tanóti* type will be dealt with in more detail in §10.42.

5.24. Reduplicated athematic presents. These are like ordinary root presents except that an additional syllable is added to the beginning that consists of the first consonant of the root plus an e or i . (Vowel-initial roots were rare in PIE, and no reduplicated presents are known to have been made from them.) For example, the root $*deh_3-$ 'give' formed a reduplicated athematic present stem $*de-deh_3-$ in the singular and $*de-dh_3-$ in the dual and plural; this is reflected in Vedic *dā-dā-ti* 'gives' and *dā-d-ati* 'they give' (the latter from PIE $*de-dh_3-nti$).

In many examples of the type from the daughter languages, the reduplicating syllable has $-i-$ rather than $-e-$, as in Vedic Skt. *jī-gā-ti* 'he goes' and Gk. *dí-dō-mi* 'I give'. This pattern probably spread from thematic reduplicated presents like Gk. *gígnomai* 'I become' (§5.33 below).

Thematic presents

5.25. Thematic presents, like other thematic formations, have a theme or stem vowel, ablauting **-e/o-*, before the personal endings. The stress was fixed and the grade of the root did not change. The personal endings of thematic presents are the same as those of athematic presents except for the 1st person singular, which was **-h₂* (or **-oh₂* when we include the thematic vowel) rather than **-mi*. This ending is ultimately the same as the 1st singular ending of the middle (**-h₂e*), and in fact it is widely believed that the thematic conjugation had its origins in the middle.

The following are sample paradigms of the thematic conjugation, showing the verb meaning 'bear, carry' in Vedic, Greek, Gothic, and Old Church Slavonic, as well as the verb meaning 'drive, do' in Latin. In the right-hand column is the reconstructed thematic present of **bher-* 'bear, carry':

	Vedic Skt.	Gk.	Lat.	Goth.	OCS	PIE
sg. 1	<i>bhārāmi</i>	<i>phérō</i>	<i>agō</i>	<i>baíra</i>	<i>berō</i>	<i>*bhér-o-h₂</i>
2	<i>bhārasi</i>	<i>phéreis</i>	<i>agis</i>	<i>baíris</i>	<i>beresi</i>	<i>*bhér-e-si</i>
3	<i>bhārati</i>	<i>phérei</i>	<i>agit</i>	<i>baíriþ</i>	<i>bereti</i>	<i>*bhér-e-ti</i>
du. 1	<i>bhārāvas</i>			<i>baíros</i>	<i>berevě</i>	<i>*bhér-o-ye-</i>
2	<i>bhārathas</i>	<i>phéreton</i>		<i>baírats</i>	<i>bereta</i>	<i>*bhér-e-to-</i>
3	<i>bhāratas</i>	<i>phéreton</i>			<i>berete</i>	<i>*bhér-e-to-</i>
pl. 1	<i>bhārāmas(i)</i>	<i>phéromen</i>	<i>agimus</i>	<i>baíram</i>	<i>beremū</i>	<i>*bhér-o-me-</i>
2	<i>bhāratha</i>	<i>phérete</i>	<i>agitis</i>	<i>baíriþ</i>	<i>berete</i>	<i>*bhér-e-te(-)</i>
3	<i>bhāranti</i>	<i>phérousi</i>	<i>agunt</i>	<i>baírand</i>	<i>berotū</i>	<i>*bhér-o-nti</i>

5.26. The theme vowel was in the o-grade before the 1st person endings and the 3rd plural, i.e. before endings beginning with a resonant or laryngeal; the reason for this is not known. It was not an infrequent occurrence in the daughter languages that the athematic 1st singular ending *-mi* was tacked on secondarily, as in Vedic *bhārāmi* and Old Church Slavonic *berō* (< **bher-ōm(i)*) above.

We next survey the major types of thematic presents.

5.27. Simple thematic presents had full grade of the root, with the accent on it, followed by the theme vowel and the ending, as **bhér-e-ti* 'bears' above. It is noteworthy that simple thematic verbs are almost or completely absent from Anatolian.

5.28. The type known as *tudāti*-presents had zero-grade of the root and accent on the theme vowel. The name comes from a representative example in Sanskrit, Vedic *tudāti* 'beats' < **tud-é-ti*.

5.29. **je/o*-presents. A thematic suffix **-je/o-* is widely represented, appearing in several different functions. A number of verbs formed their ordinary presents with this suffix, e.g. **leh₂-* 'bark' had a present stem **léh₂-je-* (as in Vedic Skt. *ráyasi* 'you bark', Lith. *lōju* 'I bark'), and to **men-* 'think' the present was **mṇ-jé-* (as in Vedic Skt. *mányate* 'thinks' and Gk. *mainetai* 'is mad' < earlier **man-je-* < **mṇ-jé-*). These are usually termed *primary *je/o*-presents (not to be confused with "primary" in the meaning "non-past"). The type with zero-grade of the root and accented suffix, characteristically used with intransitives, may have been restricted to middle inflection originally, which would explain why in some branches (Indo-Iranian, Armenian) it came to be used to form the passive.

5.30. A different but homophonous suffix, also accented, was used to form denominative verbs (verbs derived from nouns or other parts of speech, like English *to head*, *to chair*). These verbs were formed by adding the suffix **-je/o-* directly to the stem of a noun. Thus from the noun **h₁neh₃mn-* 'name' (exact preform uncertain, see §6.36) was formed a denominative verb **h₁neh₃mn-je/o-* 'to name', reflected in Hittite *lamn-ya-zzi* 'names' (from *lamn-* 'name', with *l* replacing *n*), Greek *onomaino* 'I name' (< pre-Greek **enom₁-jō*), and Germanic **namnjan* 'to name' (in e.g. Gothic *namnjan*, Old High German *nemmen*, and Modern German *nennen*).

Denominative **je/o-* verbs went on to become extremely productive in most of the daughter languages, and over time new denominative suffixes were created from them by resegmentation. For example, Greek made many denominatives from nouns whose stem ended in *-id-*; the resultant pre-Greek combination **-id-je/o-* became (by regular sound changes) Gk. *-ize/o-*, and this was then reanalyzed as a separate denominative suffix in its own right. It eventually made its way into English as the all-purpose verbal suffix *-ize*.

5.31. **skē/ō-* presents. PIE possessed an accented thematic suffix **-skē/ō-*, added to the zero-grade of the root. The productive descendants of this formation differ in meaning from branch to branch. In Anatolian, the suffix indicates repeated, habitual, or background action, or action applied to more than one object, as in Hittite *walḫ-iški-zzi* 'beats (repeatedly), beats (several objects)'. The habitual or durative sense is also found in Homeric Greek (e.g. *pheúgeskon* 'they would (habitually) flee'). Note also its use in the existential verbs Palaic *iška* and Archaic Latin *escit* 'there is'. Other Latin verbs with the suffix, however, are inchoatives (indicating the beginning or inception of an action or state), e.g. *rubē-sc-ere* 'to grow red'. Several verbs having the suffix that are reconstructible for PIE refer to asking or wishing, indicating perhaps that the suffix also once had desiderative function. An example is **p₁rk-skē-* 'ask' in Vedic Skt. *pyrchāti* 'asks', Lat. *poscit* 'asks', and German *forschen* 'to look into, research'.

5.32. Causative-iteratives. To form a verb meaning 'cause to do X', PIE took the o-grade of the root and added the accented thematic suffix **-éje/o-*. Thus the causative of **sed-* 'sit' was **sod-éje/o-* 'cause to sit' (> Old Irish [*ad-*]*suidi* 'makes sit', Gothic *satjan* 'to set, plant'); similarly, **ues-* 'clothe' has the causative **uos-éje/o-* 'clothe, put on clothes' (> Hittite *waššezzu* 'he clothes', Vedic Skt. *vāśáyati* 'he clothes'). But almost all such formations in Greek, and many in Slavic, have iterative and not causative meaning, such as Gk. *phor-éo* 'I carry around, habitually carry' (from *pherein* 'to carry') and Old Church Slavonic *nositŭ* 'habitually carries' (from *nesti* 'to carry'). This feature is probably also of PIE date, hence the term "causative-iterative" for this class of verbs.

5.33. Reduplicated thematic presents. Unlike their athematic counterparts, reduplicated thematic presents were formed with **-i-* in the reduplicating syllable and had zero-grade of the root. Some examples include Gk. *gí-gn-o-mai* 'I am born' (root **gēnh₁-* 'be born') and Lat. *sí-st-ō* 'I stand' (root **steh₂-* 'stand').

5.34. Other presents. A suffix **-h₂-*, added to thematic adjectival stems (resulting in a sequence **-eh₂-*), was used to form a verb meaning 'to make something have the quality of the adjective'. Such a verb is called a *factitive*. This factitive suffix may have been further combined with the suffix **-je/o-*. Thus from **new₁-o-* 'new' was formed the factitive **new₁-eh₂-(je)-* 'make new' in Hitt. *new-aḫḫ-* 'make new', Lat. (*rē-*)*nou-ā-re* 'to make new' (*-ā-* < **-eh₂-*).

There was also a *stative* suffix **-eh₁-* (probably also followed by **-je/o-*) added to an adjectival root to form a verb meaning 'have the quality of the adjective',

as in the cognate forms Lat. *rub-ē-re*, OHG *rot-ē-n*, and OCS *rŭdĕti*, all meaning 'to be red' from **h₁rudh-eh₁-* (from the adjectival root **h₁reudh-* 'red'). Note also Hittite *marš-e* 'be false' (< *marša-* 'false'). The suffix appears in numerous other formations in the daughter languages, such as the Greek aorist passive suffix *-ē-* (§12.43).

A new, controversial theory posits a class of presents having *o*-grade in the singular and personal endings like those of the perfect; see §9.33.

5.35. Desideratives and futures. Several suffixes containing an **-s-* have been reconstructed that were used to form desideratives, that is, verbs expressing desire or intent. Some of their descendants function as futures (see below), but it is not certain whether any of these were true futures in PIE. How, and whether, these future/desiderative formations are related to each other is still an open question.

5.36. Indo-Iranian, Balto-Slavic, and maybe Celtic point to a suffix **-sje/o-* with future and desiderative meaning: Vedic Skt. *dā-syá-ti* 'he intends/wants to give' (later 'he will give'), Lithuanian *dúosiant-* 'about to give', Old Russian *byšęši* 'future' (< **about to be*, from a participle **bhuH-sjom-*), and Gaulish *pissiumi* 'I will see'.

5.37. A reduplicated desiderative with *i*-reduplication and a suffix **(h₁)se-* is found in Indo-Iranian and Celtic. For example, from **g^hben-* 'slay' was formed **g^hhi-g^hh₁se-ti*, reflected in Vedic Skt. *jíghāṃsati* 'wants to slay' and the Old Irish future *(-)géna* 'he will slay' (from Celtic **g^hig^hmā-se-ti*). The same suffix, without reduplication, is the source of Greek futures of the type Homeric *kaléo* 'I will call' < pre-Greek **kalesō* < **kal-h₁se-*.

5.38. An athematic future suffix in **-s-* (without preceding laryngeal) is found in Italic (Umbr. *fu-s-t* 'he will be') and Baltic (Lith. *būs* 'he will be'), while Greek has a thematic version that seems to have been middle in the first instance (e.g. *dék-s-o-mai* 'I will bite' to active present *dáknō* 'I bite'); cp. §5.6.

The imperfect and injunctive

5.39. The present stem was used to form not just the present tense but also the imperfect. As mentioned above, the imperfect is usually thought to have signified durative or repeated action in past time (*was going, used to go*). Formally it was usually identical to the present except that secondary endings were used instead of primary (so **-m*, **-s*, **-t* instead of **-mi*, **-si*, **-ti* in the active). The 1st singular **-m* is the same in both thematic and athematic imperfects. The original type is best preserved in Anatolian, Indo-Iranian, and Greek: for example, Hitt. (preterite) *daškinun* 'I (repeatedly) took', Vedic Skt. *ábharam* 'I was carrying', Av. *barəm* 'I was carrying', and Gk. *épheron* 'I was carrying'. (The Vedic and Greek forms begin with the augment **e-* explained in the next section.) Full paradigms were given above in §§5.12 and 5.14.

Outside of these branches, the IE imperfect has either been completely lost, or merged with the aorist. In those branches where the imperfect was lost, a new imperfect conjugation was often innovated (as in Italic and Slavic), sometimes of obscure origin (as in Celtic).

5.40. Indo-Iranian, Greek, Armenian, and Phrygian attest a prefix called the augment that was added to past-tense forms. It is reconstructible as **e-*, as in the imperfect **é-bher-e-t* 'he was carrying' (Vedic Skt. *ábharat*, Gk. *éphere*, Arm. *eber*) or the aorist **e-dheb₁-* 'placed' (Phrygian *edaes* 'he placed'). These branches have other features in common too, and may well have emerged out of a single dialect

area of late PIE; thus it is often thought that the augment was an innovation peculiar to that dialect.

The addition of the augment was not obligatory in early Indo-Iranian and in Homeric Greek. The augmentless past-tense forms in these languages are called *injunctives*, and have been much discussed in the literature. Their precise function or functions are still not fully clear. In Homer, *injunctives* are interchangeable with past tenses but sometimes have gnomic force (that is, are used to express general truths). In Indo-Iranian, *injunctives* can indicate intent, futurity, and some quasi-modal meanings, and were also used in commands, especially prohibitions.

The Aorist Stem

5.41. PIE possessed several different aorist stem formations, some formally parallel to the present-stem types. Comparable to root athematic and simple thematic presents, for example, were root aorists and thematic aorists. In addition, PIE possessed a special aorist suffix **-s-* used to form the so-called sigmatic or *s*-aorist. Some of the aorist formations persist even in languages where there is no longer an aorist proper, such as Latin, Old Irish, and Tocharian; in these languages, such forms are called preterites or perfects. In languages with the augment (see above), the aorist is augmented. The aorist indicative was inflected using the secondary personal endings.

5.42. Root aorist. The root aorist was formed by adding the secondary endings directly to the full grade of the root in the active singular, and to the zero-grade of the root elsewhere. Thus from **steh₂-* 'stand' was formed the root aorist **(e)steh₂-t* 'he stood' (Vedic Skt. *ástthāt*, Doric Gk. *éstā*). Root aorists typically are made from roots that form characterized presents (see §5.21 on this term); Doric *éstā*, for example, is the aorist of the reduplicated present *hístāmi* 'I stand'.

5.43. Sigmatic or *s*-aorist. The PIE sigmatic or *s*-aorist was characterized by the addition of an **-s-* to the verbal root. The root is in the lengthened *e*-grade in Indo-Iranian, Italic, and Slavic, as in Vedic Skt. *(á-)vāks-ur* (< **(é-)uēgh-s-*) 'they conveyed', Lat. *uēxī* 'I conveyed' (a perfect in Latin but originally an *s*-aorist), and Old Church Slavonic *věsomŭ* (< **uēdh-s-me*) 'we led'. In Greek and in the *s*-aorist middle in Indo-Iranian, however, the root was in the full grade, as in Gk. *élekse* 'he said' (< **e-leg-s-*). The ablaut thus may have originally been of the "Narten" type (§5.20), with lengthened grade in the singular and full grade elsewhere.

The *s*-aorist eventually outstripped other kinds of aorists in Indo-Iranian, Greek, and Slavic. It is absent from Anatolian and Tocharian, where however there are preterites having an *-s* in the 3rd singular active (e.g. Hitt. *naiš* 'he brought', Toch. B *preksa* 'he asked' < **prēk-s-*). Possibly, these two branches reflect an earlier, rudimentary stage in the development of the *s*-aorist, and split off from the family before the *s*-aorist evolved further into its more familiar form.

5.44. Thematic aorist. The thematic aorist resembled the thematic imperfect, but always had zero-grade of the root. Two examples of thematic aorists reconstructible for PIE are **(é-)uid-e* 'saw' (root **uēid-* 'see': Vedic Skt. *á-vidat* 'found', Gk. *é-(w)ide* 'saw', Arm. *e-git* 'found') and **(é-)h₁ludh-e-* 'went' (root **h₁leudh-* 'go': Hom. *éluthe* 'came', Old Irish *luid*, Toch. A *lāc*).

5.45. Reduplicated aorist. On the basis of Indo-Iranian, Greek, and Tocharian A we can reconstruct a reduplicated thematic aorist for PIE. An example recon-

structible for PIE, having *e* in the reduplicating syllable and zero-grade of the root, is **ue-ukʷ-e-* 'spoke' (from the root **uekʷ-* 'speak'), seen in Vedic Skt. (*á*)*vocam* 'I spoke' and Gk. (*w*)*eĩpon* 'I spoke' (dissimilated from *(*w*)*eĩpon*). Aside from this example (and a very few others), reduplicated aorists typically have causative meaning, such as Vedic Skt. *á-pi-par-as* 'you made cross over', Gk. *dé-da-e* 'he taught' (< *'caused to know'), and Toch. A *ša-särs* 'he made known'.

5.46. Long-vowel preterites. A variety of past-tense verb forms with long root vowel are found scattered among the branches. Particularly widely represented are forms reflecting **ē* in the root. A long-vowel preterite **lēg-* 'gathered, looked at' is reflected by Latin *lēg-i* 'I gathered, read' (present *leg-ere* 'to gather'), Albanian (*mb*)*lodha* 'I gathered' (*o* < **ē* in Albanian), and Tocharian A *lyāk* 'saw'. According to one recent (and controversial) theory, these were originally imperfects of Narten presents, hence the lengthened grade. Other vowels are seen too, as in Greek *ánōge* 'he ordered', Gothic *mol* 'I/he ground' (present *mal-an* 'to grind'; the *o* is long), and Old Irish *fich* 'he fought' (present *fich-id* 'fights'). These all have diverse origins. Some, such as the Irish forms, are demonstrably more recent creations; others, such as the long-*ē* preterites, appear to be ancient. How they fit into the PIE verbal system is not known.

The Perfect Stem

5.47. The perfect stem was formed by reduplication, specifically, by doubling the first consonant of the root and inserting *e* (no perfects to vowel-initial roots in PIE are known). Characteristic of the perfect was the appearance of the root in the o-grade in the singular and accented; in the dual and plural, it was in the zero-grade and the endings were accented. Thus the perfect singular stem of **men-* 'think' was **me-món-*, and the dual and plural stem, **me-mn-*.

The perfect had a special set of personal endings that closely resemble those of the middle:

	Perfect	Primary middle	Secondary middle
sg. 1	*- <i>h₂e</i>	*- <i>h₂er</i>	*- <i>h₂e</i>
2	*- <i>th₂e</i>	*- <i>th₂er</i>	*- <i>th₂e</i>
3	*- <i>e</i>	*- <i>or</i>	*- <i>o</i>
pl. 1	*- <i>me-</i>	*- <i>medhh₂ ?</i>	*- <i>medhh₂ ?</i>
2	*- <i>e</i>	*- <i>dh(u)ue- ?</i>	*- <i>dh(u)ue- ?</i>
3	*- <i>ēr, *-rs</i>	*- <i>ro(r?)</i>	*- <i>ro</i>

The close similarity of the perfect endings to the middle endings has generated much research and controversy; precisely what the connection between them is remains unclear. Not unconnected with the formal overlap of perfect and middle endings is the fact that some perfects have middle meaning (like Gk. *ólōla* 'I am lost', vs. the active present *óllumi* 'I lose, destroy'), or exist alongside middle presents (like Vedic Skt. *ruróca* 'shines', perfect of the middle present *rócate* 'shines') or middle root aorists (Vedic Skt. perfect *jujōša* 'enjoys' next to middle root aorist *ájusran* 'they took a liking to'). It has been speculated that the perfect and the middle endings were once a single set.

The following are some of the forms upon which this reconstruction has been based. They do not all have the same function (the Hittite forms are presents and the Gothic forms are preterites), more on which below in §5.49. The forms are Hitt. *wewakḫi* 'I demand', Vedic Skt. *jagáma* 'I went', Gk. *léloipa* 'I left', Lat. *memini* 'I remember', and Goth. *haihait* 'I was called, I was named'. At the right is a reconstructed PIE paradigm.

	Hitt.	Ved.	Gk.	Lat.	Goth.	PIE
sg. 1	<i>(wewakḫi)</i>	<i>jagáma</i>	<i>léloipa</i>	<i>memini</i>	<i>haihait</i>	* <i>me-món-h₂e</i>
2	<i>(wewakti)</i>	<i>jagántha</i>	<i>léloipas</i>	<i>meministi</i>	<i>haihaist</i>	* <i>me-món-th₂e</i>
3	<i>wewakki</i>	<i>jagáma</i>	<i>léloipe</i>	<i>meminit</i>	<i>haihait</i>	* <i>me-món-e</i>
pl. 1		<i>jaganmá</i>	<i>leloípamen</i>	<i>meminimus</i>	<i>haihaitum</i>	* <i>me-mn̥-mé(-)</i>
2		<i>jagmá</i>	<i>leloípate</i>	<i>meministis</i>	<i>haihaituþ</i>	* <i>me-mn̥-é ?</i>
3		<i>jagmár</i>	<i>leloípási</i>	<i>meminēre</i> ¹	<i>haihaitun</i>	* <i>me-mn̥-ēr</i>

¹ Older or poetic; ordinary Classical *meminērunt*.

Most of the languages have undone the original ablaut and have partially replaced the perfect endings with secondary endings, especially in the plural.

5.48. One famous perfect does not exhibit reduplication, **uoid-h₂e* 'know' (from the root **uoid-* 'see'), becoming Vedic Skt. *véda*, Gk. *(w)óida*, Goth. *wait*, and Old Eng. *wāt* (continued in the modern British English phrase *God wot* 'God knows'). Compare the following paradigms:

	Ved.	Gk.	Goth.	PIE
sg. 1	<i>véda</i>	<i>(w)óida</i>	<i>wait</i>	* <i>uoid-h₂e</i>
2	<i>véttha</i>	<i>(w)óistha</i>	<i>waist</i>	* <i>uoid-th₂e</i>
3	<i>véda</i>	<i>(w)óide</i>	<i>wait</i>	* <i>uoid-e</i>
pl. 1	<i>vidmá</i>	<i>(w)ídm̐en</i>	<i>witum</i>	* <i>uid-mé(-)</i>
2	<i>vidá</i>	<i>(w)íste</i>	<i>wituþ</i>	* <i>uid-é ?</i>
3	<i>vidúr</i>	<i>(w)ísāsi</i>	<i>witun</i>	* <i>uid-ēr</i>

This perfect is exceptional, not only because of its lack of reduplication but also because of its meaning (not 'sees' or 'has seen' but 'knows'). These are widely thought to be archaic features, harking back, perhaps, to a time when the perfect had no reduplication; but this view is not universally held.

Meaning of the perfect

5.49. The perfect is formally best preserved in Indo-Iranian and Greek, and less well in Anatolian, Italic, Celtic, and Germanic. A number of archaic examples of the perfect, especially in Indo-Iranian and Greek, refer to states in present time. We saw this above (§5.47) in four perfects: Gk. *ólōla* 'I am lost', Vedic Skt. *ruṛóca* 'shines', Ved. *jujōsa* 'enjoys', and Lat. *meminit* 'remembers'. (The Greek and Gothic cognates of *meminit* are also statives, meaning 'is mindful of': Gk. *mémone*, Goth. *man* 'thinks' [without reduplication].) Furthermore, the singular perfect endings are used to inflect a class of presents in Anatolian, the so-called *hi*-conjugation (to which Hittite *wewakki* in §5.47 belongs). While the relationship of the *hi*-conjugation to

the perfect is unclear, all these facts together have led researchers to believe that the PIE perfect was a *stative*. In the daughter languages, however, except for relic forms like the ones just cited, perfects express past tense, and have often fallen together formally with the aorist into a single “preterite” tense. To explain this development, it is usually said that the PIE stative perfect had (or optionally had) resultative overtones (‘is in a state resulting from having done X’, therefore ‘has done X’).

Moods

Imperative

5.50. The imperative was used to express direct commands. The athematic 2nd singular imperative ending is reconstructible as **-dhi* and was added to the zero-grade of the root, as in Vedic Skt. *śru-dhí* ‘listen!’ and Gk. *í-thi* ‘go!’. The thematic 2nd singular imperative was the bare thematic stem, as in Vedic Skt. *bhára* and Gk. *phére* ‘carry!’ from PIE **bhér-e*. The 2nd plural imperative ended in **-te*: Vedic Skt. *bhárata* and Gk. *phérete*, both from **bhérete* ‘(you pl.) carry!’

PIE also had 3rd person imperatives ending in **-u*, forming 3rd sing. **-tu* and 3rd pl. **-ntu*, as in Hittite *paiddu* ‘let him go’ and *pāndu* ‘let them go’. In Greek, Italic, and Celtic, however, the 3rd person imperative endings go back to **-tōd* and **-ntōd*, as in Greek *phéré-tō* ‘let him carry, he shall carry’, Archaic Latin *da-tōd* ‘he shall give’, and Celtiberian *tatuz* (phonetically probably *datuz*) ‘let him give’.

Subjunctive

5.51. The subjunctive in PIE was formed by the addition of the thematic vowel to the verb stem (be it athematic or already thematic), followed apparently by primary endings (although in Indo-Iranian both primary and secondary endings were used). In athematic verbs, the “strong” stem (the one having full grade of the root or ablauting portion of the stem) was used in all persons and numbers. Some reconstructed examples of 3rd sing. and 3rd pl. present subjunctives follow for the athematic presents of **h₁es-* ‘be’ and **jeug-* ‘yoke’ and the thematic present **bher-* ‘carry’:

	Indicative	Subjunctive
root present	<i>*h₁és-ti</i> <i>*h₁s-énti</i>	<i>*h₁és-e-ti</i> <i>*h₁és-o-nti</i>
nasal-infix present	<i>*ju-né-g-ti</i> <i>*ju-n-g-énti</i>	<i>*ju-né-g-e-ti</i> <i>*ju-né-g-o-nti</i>
thematic present	<i>*bhér-e-ti</i> <i>*bhér-o-nti</i>	<i>*bhér-e-e-ti</i> (= <i>*bhérēti</i>) <i>*bhér-o-o-nti</i> (= <i>*bhérōnti</i>)

The right-hand column above yields the Vedic subjunctives *ásat(i)*, *ásanti* (versus indicative *ásti*, *sánti*); *yunájati*, *yunájanti* (versus indicative *yunákti*, *yunjánti*); and *bhárati*, *bhárānti* (versus indicative *bhárati*, *bháranti*). Compare also Latin *erit* ‘he will be’, *erunt* ‘they will be’ (synchronically future, but historically subjunctive), and Greek *phérēi*, *phérōsi* for the thematic subjunctive.

As indicated by such first-person subjunctives as Vedic Skt. *kṛṇavā* 'I will do', Old Avestan *yaojā* 'I will yoke', Gk. *phérō* 'let me carry', and Lat. *erō* 'I will be', the 1st singular ended in **-h₂* (or **-oh₂* including the subjunctive vowel) rather than **-mi*.

Subjunctives could also be formed in the same way from root and s-aorists, where likewise the full grade of the aorist stem was used.

5.52. Meaning of the subjunctive. The subjunctive was probably a future tense. Here one must be careful to avoid terminological confusion, as most of the forms called "subjunctives" in the daughter languages, which express a variety of modal meanings, have nothing to do with the PIE subjunctive, but come from the optative (see below) or elsewhere. The subjunctive in Indo-Iranian, Greek, and partly in Celtic is a continuation of the PIE subjunctive; in Indo-Iranian it usually has future meaning, as true also of many examples in Homer. In Latin, the descendants of the PIE subjunctive are futures. The PIE subjunctive is not preserved outside of these branches.

Optative

5.53. The optative was formed by adding an ablauting suffix **-ieh₁-/-ih₁-* to the relevant stem (present, aorist, or perfect), plus secondary personal endings. In most athematic verbs, this suffix was attached to the weak stem of the verb, and had the full-grade form **-ieh₁-* in the singular and zero-grade **-ih₁-* elsewhere. However, in Narten presents, being accented on the root, the suffix was in the zero-grade throughout. This was also true of thematic verbs, where the zero-grade of the suffix was used throughout, and added to the o-grade of the thematic vowel. Compare the following paradigms, showing the optatives of the athematic verb 'be' and the thematic verb 'bear' in PIE and selected daughter languages:

Athematic optative					Thematic optative			
	Ved.	Gk.	Lat.	PIE	Ved.	Av.	Gk.	PIE
sg. 1	<i>syām</i>	<i>eîēn</i>	<i>sem</i> ¹	<i>*h₂s-ieh₁-m</i>	<i>bhāreyam</i>		<i>phéroumi</i> ²	<i>*bhér-o-ih₁-m</i>
2	<i>syās</i>	<i>eîēs</i>	<i>siēs</i> ¹	<i>*h₂s-ieh₁-s</i>	<i>bhāres</i>	<i>barōiš</i>	<i>phérouis</i>	<i>*bhér-o-ih₁-s</i>
3	<i>syāt</i>	<i>eîē</i>	<i>siēt</i> ¹	<i>*h₂s-ieh₁-t</i>	<i>bhāret</i>	<i>barōišt</i>	<i>phérois</i>	<i>*bhér-o-ih₁-t</i>
pl. 1	<i>syāma</i>	<i>eîimen</i>	<i>simus</i>	<i>*h₂s-ih₁-mé-</i>	<i>bhārema</i>	<i>baraēma</i> ¹	<i>phéroumen</i>	<i>*bhér-o-ih₁-me-</i>
2	<i>syāta</i>	<i>eîite</i>	<i>sitis</i>	<i>*h₂s-ih₁-té-</i>	<i>bhāreta</i>		<i>phéroute</i>	<i>*bhér-o-ih₁-te-</i>
3	<i>syūr</i>	<i>eîen</i>	<i>sient</i> ¹	<i>*h₂s-ih₁-ént</i>	<i>bhāreyur</i>	<i>barauan</i>	<i>phérouen</i>	<i>*bhér-o-ih₁-ent</i>

¹ Archaic Latin; Classical *sim*, *sīs*, *sit*, 3rd pl. *sint*. ² The *-mi* is an innovation; cp. Arcadian *ekselaunoia* 'I would drive out', with *-oia* from **-oih₁-m*. ³ This form not actually attested for this verb.

As mentioned above, the continuations of the PIE optative in the daughter languages are not always termed "optatives," but sometimes "subjunctives," as in Italic and Germanic. In Balto-Slavic the optative survives as an imperative or a form called the permissive. There is no trace of the optative in Anatolian.

Non-finite Verbal Formations and Other Topics

We have now concluded our survey of the finite forms of the IE verb, that is, those that differentiate person and number. Two other sets of forms remain: infinitives and participles.

Infinitives

5.54. The infinitive is essentially a verbal noun, rendered either as 'to X' or 'X-ing' in English. The daughter languages exhibit a rather bewildering variety of infinitives. Typically they are frozen case-forms (usually accusatives, datives, or locatives) of nouns derived from verbal roots. The nominal formations in question are usually old, but which infinitive formations are of PIE date is uncertain. Represented in more than one daughter branch as true infinitives are the following:

- 1 The suffix **-dhie-* or **-dhjo-*, appearing in both active and passive infinitives: Vedic Skt. *pība-dhyai* 'to drink', Umbr. *pīha-fi* 'to propitiate', Toch. A and B *lkā-tsi* 'to look'.
- 2 Various case-forms of the noun suffix **-tu-* (§6.42), e.g. Vedic Skt. *dā-tum* 'to give', *pā-tave* 'to drink, for drinking'.
- 3 Various case-forms of the noun suffix **-ti-* (§6.42), e.g. Vedic Skt. *pī-tāye* 'to drink, for drinking', Av. *kərə-tāe* 'to do', OCS *da-ti* 'to give', Lith. *bū-ti* 'to be'.
- 4 Various case-forms of the complex *n*-stems (§6.34) **-men-*, **-sen-*, **-ten-*, and **-uen-*, e.g. Vedic Skt. *vid-māne* 'to find', Homeric Gk. *id-menai* 'to know', the Gk. thematic infinitive ending *-ein* (< **-e-sen*), OPers. *car-tanaīy* 'to do', Hitt. *lahhiya-uwanzi* 'to wage war', and Cypriot Gk. *do-wenai* 'to give'.

5.55. Several daughter branches have a specific infinitive formation often called the *supine* that is solely used with verbs of motion to indicate purpose, as Vedic (*āganma*) *prāvohum* '(we went) to bring', Latin (*uenērunt*) *questum* '(they came) to complain', Old Church Slavonic (*pridŭ*) *sŭpatŭ* 'I went (in order) to sleep'. Usually it is formed with the suffix **-tum*, the accusative of the abstract noun suffix **-tu-* mentioned above.

Participles

5.56. Participles are verbal adjectives. Like other adjectives, they inflect in the different nominal cases and numbers discussed in chapter 6.

The most widely represented participle is the *nt*-participle, found in virtually all the branches to form participles of active voice to present or aorist stems. The suffix was ablauting, appearing as **-ent-* (perhaps also as **-ont-*) and **-nt-*. In athematic verbs, it was added to the weak present stem, as in **h₃s-(e)nt-* 'being' (> Vedic Skt. *sat-* 'being', Lat. *ab-sent-* 'being away, absent'), **dhe-dbh₁-(e)nt-* 'placing' (> Gk. *ti-thent-* 'placing'), and **jung-(e)nt-* 'joining' (> Lat. *iungent-*). In thematic verbs, zero-grade **-nt-* was added to the o-grade thematic vowel, as in **bher-o-nt-* 'carrying' (> Gk. *phéront-*, Goth. *baírand-*, and OCS *berǫst-* from pre-Slavic **ber-ont-j-*). In Anatolian, however, this participle is semantically equivalent to the **tó*-verbal adjective (see §5.57 below), indicating completion and being passive when formed from transitive verbs, and active when formed from intransitive verbs: *kunānt-* '(having been) killed', *iyant-* '(having) gone'. Aorist active participles were formed similarly to present participles, as the root aorist participle meaning 'having stood' in Vedic Skt. *sthānt-* and Gk. *stant-*.

Several branches continue a mediopassive participle in **-m(e)no-* or **-mh₃no-* (its exact reconstruction is a matter of dispute), seen for example in Vedic Skt. *bhāra-māna-* and Av. *bara-mna-* 'carrying (for oneself)', Gk. *phéró-menos* 'carrying

(oneself), being carried', and OPruss. *poklausīmanas* 'heard'. In other branches only a few fossilized examples of it exist, such as Lat. *alu-mnus* 'fosterling' (lit. 'nurtured') and Arm. *anasoun* 'animal' (lit. 'not talking', from *asem* 'I talk', as though from late PIE **n-aḡ-omno-*). Related is the present passive participle in **-mo-* of Balto-Slavic and perhaps Anatolian: OCS *něsomŭ* 'being carried', Lith. *nėšamas* 'being carried', and Cuneiform Luv. *kīšammi-* 'combed'.

A perfect participle with the ablauting suffix **-uos/-us-* is reflected in numerous branches. In PIE, it was added to the zero-grade of the perfect stem; the suffix was **-uōs* in the nominative singular masculine (neuter **-us*, feminine **-us-ih₂*); the accusative singular ended in **-uos-m* (see §6.6 for the case-ending) and the other cases were added to the stem **-us-*. This is directly reflected for example in Av. nomin. sing. *viduua*, accus. *viduuaighəm*, stem *viduś-* 'knowing' (from **uiduōs* **uiduōsm* **uidus-*). Compare also the Vedic stem *vidūś-* 'knowing'; Gk. masc. nomin. sing. *(w)eidōs* and fem. *(w)idūia* (< **uidusih₂*) 'knowing'; and, further afield, Toch. B masc. nomin. *lt-u* 'having come out', accus. *lt-uwes*, fem. *lt-usa*, and Lith. fem. *aug-us-i* 'having grown'.

Some branches also have evidence for a preterite participle in **-lo-*: Arm. aorist participle *gereal* 'taken, having taken', OCS preterite participle *nes-lŭ* 'having carried', and Toch. A gerundive (verbal adjective) *ritwāl* 'united'.

5.57. In addition to the participles, PIE had verbal adjectives in **-tō-* and **-nó-*, added to the zero-grade of a verbal root. These indicated completed action and were semantically like past participles in English: if the verb they were formed from was transitive (like *eat*), the verbal adjective was passive and past in tense (*eaten*), but if the verb was intransitive (like *go*), the verbal adjective was simply past in tense (*gone*). Thus the root **g^hen-* 'kill' formed the verbal adjective **g^hen-tō-* 'slain' in Vedic Skt. *hatá-* and Gk. *(-)phatós*, and the intransitive root **g^hem-* 'come' has **g^hem-tō-* '(having) come' in Vedic Skt. *gatá-*, Gk. *(-)batós*, and Lat. *uentus*. Less widespread than **-tō-* was **-nó-*, whose participial function is clearest in Indo-Iranian, as in Vedic Skt. *bhin-ná-* '(having been) split' (< **bhid-nó-*), but also in Germanic and Slavic in a longer form **-enó-* or **-onó-*, as in Goth. *bit-ans*, Eng. *bitt-en* (< **bhid-onó-*) and OCS *nes-enŭ* 'carried'. Similar forms, like **p^hh-nó-* 'full', were once participles ('filled up') but have descendants that are only adjectives and no longer part of any verbal paradigm: Vedic Skt. *pūrṇá-*, Lat. *plēnus* (versus past participle *-plētus* 'filled'), Goth. *fulls* (double *-ll-* < **-ln-*), Old Irish *lán*, and Lith. *pilnas*.

Neither of these formations is found in Anatolian. For other uses of **-tō-*, see §6.77.

Verbal composition

5.58. Verbs were often combined with adverbs to modify their meaning. Such adverbs were called *preverbs* and in the first instance remained separate words. Over time they tended to join with verbs as *prefixes*. As there are some interesting syntactic phenomena associated with this subject, we will defer discussion of preverbs until §§8.9ff.

Occasionally, verbs were compounded with a non-adverbial element, such as a noun. The most familiar example of this is the verb for 'believe, trust', **k^hred d^heh₁-*, literally 'place one's heart in', which became Vedic Skt. *śrād dadhāti* 'trusts, believes', Lat. *credō* 'I believe', and Old Irish *cretim* 'I believe'.

Prosodic status of verbs

5.59. There is good comparative evidence that finite verbs were prosodically weaker (that is, were pronounced with weaker stress or lower pitch) than other parts of speech, especially in main clauses. In Vedic Sanskrit, main-clause finite verbs that do not stand at the beginning of their clause are written in the manuscripts without accent marks. In Greek, the rules for accenting verbs are different from those for nouns, and resemble the accentuation of strings of clitics; this suggests an affinity between the prosody of verbs and the prosody of chains of weakly stressed or unstressed particles. In Germanic heroic poetry, fully stressed words alliterate with one another, but certain verbs, together with unstressed pronouns and particles, do not participate in alliteration; this suggests weaker prosodic status for those verbs. In certain Germanic languages, such as modern German, verbs are required to be the second syntactic unit in main clauses, which is the same position taken by many unstressed sentence particles elsewhere in Indo-European (Wackernagel's Law, see §§8.22ff.).

All these facts taken together suggest that finite verbs were in some way prosodically deficient in PIE. Whether verbs in PIE were true clitics, that is, had no stress and formed an accentual unit with a neighboring stressed word, is uncertain, but is a position defended by most Indo-Europeanists. However, it is clear that even if they could behave as clitics some of the time, they were fully stressed when moved to the front of a clause for emphasis or contrast, or when occurring in subordinate clauses. This is not too surprising, since weaker prosodic status of verbs (*vis-à-vis* nouns) is a common cross-linguistic phenomenon.

For Further Reading

Comprehensive recent works on the verb are wanting; the volumes on the verb in Brugmann 1897–1916 and Hirt 1927–37 (see Bibliography, ch. 1) are rather out of date. Books on the IE verb are generally specialized. Watkins 1969 is a very influential examination of the prehistory of PIE verbal inflection. Jasanoff 1978 is an in-depth study of the relationship between the stative and middle; see also his most recent work, Jasanoff 2003, a very thought-provoking and detailed theory on the prehistory of IE verbal morphology. Very useful and up-to-date – though in various places controversial – is Rix 2001, a dictionary of Indo-European verbal roots and reconstructed stem-forms.

For Review

Know the meaning or significance of the following:

tense-stem	mediopassive	primary endings	<i>tudāti</i> -present
primary verb	mood	secondary endings	* <i>je/o</i> -present
derived verb	indicative	"hic et nunc" particle	augment
voice	imperative	root present	injunctive
active	subjunctive	Narten present	sigmatic aorist
middle	optative	nasal infix	

Exercises

- 1 Name the function(s) of the following PIE suffixes:

a * <i>-nt-</i>	d * <i>-eh₂-(i)e-</i>	g * <i>-eh₁-(i)e-</i>	i * <i>-ie/o-</i>	m * <i>-ieh₁-/-ih₁-</i>
b * <i>-nó-</i>	e * <i>-ne-</i> (infix)	h * <i>-e-</i> (prefix)	k * <i>-éje-</i>	n * <i>-ské/o-</i>
c * <i>-uos/-us-</i>	f * <i>-tó-</i>	l * <i>-sie/o-</i>	l * <i>-m(e)no-</i> or * <i>-mh₁no-</i>	

- 2 Identify the following reconstructed PIE verb inflectional endings as specifically as possible:

a * <i>-th₂er</i>	e * <i>-ye-</i>	l * <i>-s</i>	m * <i>-mi</i>	q * <i>-dhi</i>
b * <i>-oh₂</i>	f * <i>-(i)or</i>	j * <i>-(é)nti</i>	n * <i>-ntu</i>	r * <i>-t</i>
c * <i>-ti</i>	g * <i>-m</i>	k * <i>-tu</i>	o * <i>-ro</i>	s * <i>-h₂er</i>
d * <i>-(i)o</i>	h * <i>-si</i>	l * <i>-te</i>	p * <i>-(é)nt</i>	t * <i>-ér, *-rs</i>

In exercises 3–8, indicate the position of the accent in each PIE form that you give.

- 3 The following verbal roots formed root presents of the ordinary (non-Narten) type in PIE. Provide all three singular forms and the 3rd plural in the present indicative active.

a * <i>k₁tei-</i> 'settle'	c * <i>syen_h₂-</i> 'make a sound'	e * <i>g^when-</i> 'slay'
b * <i>h₂ueh₁-</i> 'blow'	d * <i>h₁ei-</i> 'go'	f * <i>b^hieh₁-</i> 'weep'

- 4 The following verbal roots formed simple thematic presents. For each one, provide (1) the 1st and 2nd singular and 3rd plural present indicative active, and (2) the 3rd singular and 3rd plural imperfect indicative active. Include the augment where possible.

a * <i>pek^w-</i> 'cook'	c * <i>bheidh-</i> 'trust'	e * <i>leg^w-</i> 'gather'
b * <i>pleu-</i> 'swim'	d * <i>der-</i> 'tear, flay'	f * <i>dheg^wh-</i> 'burn'

- 5 The following verbal roots formed nasal-infix presents. For each one, provide (1) all three singular forms and the 3rd plural in the present subjunctive active, and (2) the present optative stems.

a * <i>bheid-</i> 'split'	c * <i>k₁leu-</i> 'hear'	e * <i>terd-</i> 'bore'
b * <i>peuh₂-</i> 'cleanse'	d * <i>k^wreih₂-</i> 'exchange'	f * <i>h₁meigh^w-</i> 'urinate'

- 6 The following roots formed perfects. Provide all three singular forms and the 3rd plural for each perfect.

a * <i>h₂nek^w-</i> 'reach, attain'	c * <i>d₁hers-</i> 'dare'	e * <i>neig^w-</i> 'wash'
b * <i>duei-</i> 'fear'	d * <i>gen_h₁-</i> 'beget'	f * <i>g^wem-</i> 'come, go'

- 7 Form the causative-iterative stem for each of the following verbal roots and give a translation.

a * <i>b^heudh-</i> 'wake'	d * <i>leg^w-</i> 'lie'
b * <i>leuk-</i> 'light'	e * <i>dem_h₂-</i> 'tame'
c * <i>ieudh-</i> 'fight'	f * <i>nek^w-</i> 'disappear, come to harm'

- 8 The following are a mixture of thematic and strong athematic singular present stems. For each stem, (1) identify as specifically as possible the type of present it exemplifies, and (2) provide the singular and plural optative stems.

a * <i>g̑h₃-ské-</i> 'know'	d * <i>li-né-kʷ-</i> 'leave'	g * <i>sél-</i> 'jump'
b * <i>dék-</i> 'receive'	e * <i>gʷh₃-é-</i> 'devour'	h * <i>si-sd-e-</i> 'seat'
c * <i>pérd-e-</i> 'fart'	f * <i>m₁-jé-</i> 'disappear, die'	i * <i>dhe-dheh₁-</i> 'put'

- 9 The following roots formed root aorists in PIE. Provide the 2nd singular indicative active for each. Include the augment.

a * <i>peh₃-</i> 'drink'	c * <i>dreh₂-</i> 'run'	e * <i>kʷer-</i> 'cut'
b * <i>derk-</i> 'see'	d * <i>kʷel-</i> 'turn'	f * <i>gēnh₁-</i> 'beget'

- 10 The following roots formed *s*-aorists in PIE. Provide the 3rd singular indicative active for each. Include the augment.

a * <i>h₃neid-</i> 'blame, scorn'	c * <i>neih-</i> 'lead'	e * <i>bher-</i> 'carry'
b * <i>deuk-</i> 'lead'	d * <i>h₃reǵ-</i> 'direct, rule'	f * <i>prek-</i> 'ask'

- 11 Based on the endings given in §5.14 and the ablaut information in §5.19, provide the singular and the 3rd plural present and imperfect indicative middle of the verb **meuH-* 'speak'.
- 12 For 3a–c and 4a–c, give the stem of the present active participle.
- 13 For 6a–c, give the nominative singular of the perfect participle.
- 14 Using the suffix *-tó-, form the verbal adjectives of the roots in 9 above and translate them.
- 15 Given that *-ss- was simplified to *-s- in PIE (§3.37), how would you explain the Homeric form *essi* 'you (sing.) are', which looks as if it should come from **hes-si*? Do not assume any additional sound changes.
- 16 Vedic Sanskrit has an athematic verbal stem *duh-* 'to milk' (< PIE **dhugh-*, zero-grade of **dheugh-*) for which several variant forms are attested for the 3rd pl. present indicative middle, including *duháte*, *duhré*, and *duhráte*. In *duháte* and *duhráte*, the *a* is the Sanskrit outcome of **ā*, and *e* comes from **oi*. Using §5.15 as a starting point, provide historical explanations for all three forms. Which do you think is the oldest?
- 17 No IE language preserves the Narten type of root present completely intact. It has been posited on the basis of such athematic present forms as the following: Lat. 2nd sing. *ēs* 'you eat' and 3rd sing. *ēs-t* 'he/she eats'; Vedic Skt. 3rd pl. *ad-ánti* 'they eat' (*ad-* < **h₁ed-*); Vedic Skt. 3rd sing. *máś-ṭi* 'wipes' (*-ā-* < **-ē-*); Gk. 3rd sing. middle *hēs-tai* 'puts on clothes' (*hes-* < **ues-*) and *keī-tai* 'lies (down)'. Discuss how these forms support the reconstruction of the Narten type of root present.

6 The Noun

Introduction

Case

6.1. Like verbs, PIE nouns were highly inflected. Nouns had case-endings that indicated their grammatical function, such as subject, direct object, indirect object, and possessive. Sure evidence exists for eight cases in PIE: nominative (subject of the sentence and predicate nominative), vocative (the case of direct address), accusative (direct object), genitive (possessive), ablative (source or place from which), dative (indirect object, possession, and beneficiary of an action), instrumental (means and agent), and locative (place where). There may have also been a ninth case, the directive or allative (place to which). A few languages, such as Old Lithuanian, Tocharian, and Ossetic, developed additional cases that arose under the influence of neighboring, unrelated language groups (Finnic in the case of the first, Turkic in the case of the second, and Kartvelian in the case of the third).

Nominal declensions, like verbal conjugations, could be athematic or thematic. The latter gradually replaced the former in most branches, though not as quickly or thoroughly as thematic verbs replaced athematic verbs.

Number

6.2. Nouns were inflected in three numbers, singular, dual, and plural. The dual has tended to disappear, and several branches had lost it entirely or almost entirely by the time of their earliest attestations. Existing perhaps outside the singular-plural parameter was a category called the collective, which indicated a collection of entities treated as a unit (e.g., Latin *loca* '(group of) places', as opposed to the ordinary plural *loci*, singular *locus*). The collectives are only preserved as such in the archaic stages of a few branches, and were reanalyzed as singulars or plurals elsewhere. Interestingly, both Tocharian and Anatolian could form plurals to collectives, called pluratives; whether this is an inherited feature is not known, for it could have developed independently (as it later did in Celtic; cp. §14.66).

Gender

6.3. Most of the older IE languages show a three-way contrast in grammatical gender between masculine, feminine, and neuter. But the oldest preserved branch, Anatolian,

has only a two-way distinction between **animate or common gender** and **inanimate or neuter**; the historical status of the **feminine** in Anatolian is disputed.

The neuter had the same endings as the **animate** (or **masculine**) except in the **nominative**, **vocative**, and **accusative**. These three cases are not formally distinguished from one another in neuters.

Athematic Nouns

6.4. Like **athematic verbs**, **athematic nouns** have **inflectional endings** added directly to a root or suffix without an intervening **thematic vowel**. Traditionally, those **athematic nouns** with stems ending in a **consonant** (called *consonant stems*) have been distinguished from those ending in *i* or *u* (the *i*-stems and *u*-stems); but this distinction is both unnecessary and misleading, as it masks the fundamentally identical behavior of both groups over against that of the **thematic nouns**.

To understand **athematic nominal inflection**, one must distinguish between the so-called **strong** and **weak** cases. The **strong** cases differ from the **weak** cases typically in where the accent is located and which **morpheme** is in the **full grade**; most commonly, the **full grade** and the **accent** shift rightward in the **weak** cases, comparable to the shift seen in most **athematic verbs**. In the schematic diagram below, the **strong** cases are shaded:

	singular	dual	plural
nominative			
vocative			
accusative			
genitive			
ablative			
dative			
instrumental			
locative			

According to some theories, the **accusative plural** was **strong** as well. Standing somewhat apart was the **ablaut** of the **locative singular**; see §6.11.

Case-endings of athematic nouns

6.5. The reconstructions of the **PIE case-endings** are most secure for the **singular** and for the **strong** cases in the **dual** and **plural**. For the **nominative** and **accusative**, a distinction must be maintained between **animate nouns** (**masculine** or **feminine**) and **neuters** (or **inanimates**), as will become clear below.

First, we sketch the case-endings of representative daughter languages. The vocative has been omitted since the daughter languages usually used the nominative for the vocative (but see further below). Only those forms relevant to the reconstructions given later are shown.

THE SINGULAR

Representative animate athematic paradigms are given in the chart below: Hitt. *hūmant-* 'every', Ved. *pad-* and Gk. *pod-* 'foot', Ved. *gopā-* 'cowherd, protector', Gk. *klōp-* 'thief', Av. *kahrp-* 'body' and *nar-* 'man', and Lat. *uōc-* 'voice'. Baltic has clear cognate case-endings but not in any one paradigm, so a variety of forms are given that are glossed in the notes below. Case-forms in the individual languages that are either unattested or not useful for reconstruction are left blank in this and the following charts. No separate ablative is shown; see §6.8.

	Hitt.	Ved.	Av.	Gk.	Lat.	Baltic	PIE
nom.	<i>hūman-za</i>	<i>pād,</i> <i>gopā-s</i>	<i>kərəf-š</i>	<i>poús,</i> <i>klōp-s</i>	<i>uōx</i> (= <i>uōc-s</i>)	<i>viešpat-s</i> ³	*-s
acc.	<i>hūmant-an</i>	<i>pād-am,</i> <i>gopā-m</i>	<i>kahrp-am</i>	<i>pód-a</i>	<i>uōc-em</i>	<i>šūn-i</i> , ⁴ <i>smument-in</i> ⁵	*-m
gen.	<i>hūmant-aš</i>	<i>pad-ás</i>	<i>kahrp-ō</i>	<i>pod-ós</i>	<i>uōc-is</i>	<i>šūn-s</i> , ⁴ <i>szird-es</i> ⁶	*-(é)s
dat.		<i>pad-é</i>	<i>nair-e</i>		<i>uōc-i</i>	<i>giwānt-ei</i> ⁷	*-ei
instr.		<i>pad-ā</i>	<i>kahrp-a</i>				*-(e)h ₁
loc.	<i>hūmant-i</i>	<i>pad-i</i>	<i>nair-i</i>	<i>pod-i</i> ¹	<i>uōc-e</i> ²	<i>wieszpati-p</i> ⁸	*-i

¹ Synchronically labeled dative: see §6.11. ² Synchronically labeled ablative: see §6.11.
³ OLith., 'lord'. ⁴ Lith., 'dog'. ⁵ OPruss., 'man'. ⁶ OLith., 'of the heart'. ⁷ OPruss., 'for (one) living'. ⁸ OLith. adessive case, 'at the lord'.

6.6. In athematic nouns, the animate **nominative singular** ended in *-s and the **animate accusative singular** ended in *-m (or *-m after consonants). The **neuter nominative-accusative singular** had zero ending. Thus contrast Hittite animate nomin. sing. *huppar-š* (written *hupparaš*) 'bowl' with neuter *uttar* 'word' (*uttar-ø*).

Athematic animate nouns ending in resonants show lengthened grade of the stem and lack *-s, as for example Greek *patér* 'father'. It is widely held that these nominative singulars originally had ordinary full grade plus *-s also, and that a sound change (Szemerényi's Law, §3.38) happened in PIE whereby the -s was lost with compensatory lengthening of the vowel: **ph₂t-ers* > **ph₂t-ēr*. A few nouns, notably the words for 'sky, sky-god' (**d₁ēus*) and 'cow' (**g^wōus*), apparently had both lengthened grade and *-s.

6.7. In athematic animate nouns, the vocative ending was zero and there was retraction of the accent. What this meant for all practical intents and purposes is that the vocative was the same as the nominative without the *-s or without lengthened grade: Gk. nomin. *pólis* 'city' but voc. *póli* '(o) city'; Ved. nomin. *hastī* 'having a spear' (< **hastín*), voc. *hástin*. Vedic Sanskrit best preserves the accent retraction.

6.8. The **genitive** and **ablative** of athematic nouns were not distinguished from one another in PIE in the singular; they are often referred to together as the **genitive-ablative**. The ending was *-és when accented, otherwise zero-grade *-s. The first survives in genitives like Archaic Lat. *Vener-es* 'of Venus' (Classical Lat. *Veneris*),

ON *feðr* 'of a father' (< pre-Germanic **patr-es*), and Lith. dialectal *dukter-ės* 'of a daughter'. Zero-grade **-s* is found for example in Hitt. *nekuz* (pronounced *nek^uts*) 'of the evening' (< **nek^ut-s*) and OE *brōþor* 'of the brother' (< **bhrāt^r-s*). Some languages have a genitive going back to **-os*, spread perhaps from the thematic declension (§6.48): Homeric Gk. *gēne-os* 'of a kind' and Ogam Irish *Lugudecc-as* 'of Lugaid' (personal name).

6.9. The dative ended in **-ei*, directly preserved for example in Archaic Lat. *Mārt-ei* 'for Mars', Oscan *Mamert-ei* 'for Mars', and OPruss. *giwānt-ei* 'for one living'.

6.10. The instrumental ending was **-h₁* or **-eh₁*. The bare laryngeal is found in archaic Vedic *i-* and *u-*stem instrumentals in *-ī* and *-ū* from **-i-h₁* and **-u-h₁*, e.g. *matī* 'with thought'. The fuller form **-eh₁* is found in Indo-Iranian, e.g. in *vācā* 'with speech' and perhaps in the first member of such Latin verbs as *rubē-facere* 'to redden', if this originally meant 'to make with redness' or the like.

6.11. Two types of locative are found. The first was formed with a suffix **-i*, as in Ved. *pad-i* 'on the foot', Lat. *ped-e* (synchronically called an ablative, but functionally instrumental, ablative, and locative) 'with/on the foot', and Gk. *pod-i* 'for the foot' (synchronically called a dative, but functionally dative, instrumental, and locative).

The second type is the **endingless locative**, which is interesting not only in being endingless, but in typically having full or lengthened grade of the stem, in contrast to the other weak cases. For example, the Vedic endingless locative of the word for 'name' is *nām-an*, whereas the weak stem is *nām-n-* with zero-grade of the *n*-stem suffix (see below on *n*-stems). Compare also Old Avestan *dām* 'in the house' (< **dōm* or **dēm*) and Old Irish *talam* '(on the) earth' (< Common Celtic **talamon*, also an *n*-stem as in Vedic).

6.12. A ninth case, the **directive** (or allative), is posited by some on the basis primarily of Anatolian, the only branch in which it is still productive, at least in Old Hittite (e.g. *arun-ā* 'to the sea', *parn-ā* 'to the house, homeward', with long *-ā* probably indicating stress; see §9.24). Elsewhere in IE there may be a few fossilized traces extended by the locative ending **-i*, such as Gk. *khamaí* 'to the ground' < **dhǵh₁m-a*. Its PIE shape is uncertain; candidates include **-h₂(e)*, **-(e)h₂*, or simply **-a*.

THE DUAL

6.13. Fewer cases were distinguished in the dual than in the singular. The **nominative**, **vocative**, and **accusative dual** ended in **-h₁* or **-h₁e* in animate nouns, and in **-ih₁* in neuters. The full animate form is reflected for example in OLith. *žmun-e* 'two men', while the neuter is found in Gk. *ósse* and OCS *oči* 'both eyes' (**h₃(e)k^wi-h₁* and **h₃ek^w-i-h₁*, respectively, with different syllabification of the ending in the two languages), as well as Vedic Skt. *urv-ī* 'the two earths'. Note also the number 'two' itself, **d(u)uoh₁* (becoming **d(u)uō*).

The other cases of the dual cannot be reconstructed because the paradigms of the daughter languages differ too sharply from one another.

THE PLURAL

The cases of the plural also differ from one another across the daughter languages, especially outside the nominative, vocative, and accusative, making exact reconstruction uncertain in several instances. Below are some representative animate athematic

plural paradigms for comparison: Hitt. *hūmant-* 'all', Ved. *pad-* and Gk. *pod-* 'foot', Av. *kəhrp-* 'body' (and other nouns glossed in the notes), Lat. *uōc-* 'voice', and Lith. *šun-* 'dog'.

	Hitt.	Ved.	Av.	Gk.	Lat.	Lith.	PIE
nom.	<i>hūmant-eš</i>	<i>pād-as</i>	<i>kəhrp-as</i> ¹	<i>pód-es</i>		<i>šun-es</i>	*-es
acc.	<i>hūmant-uš</i>	<i>pad-ás</i>	<i>zam-as</i> ²	<i>pód-as</i>		<i>šun-is</i>	*-ns
gen.	<i>hūmant-an</i>	<i>pad-ām</i>	<i>kəhrp-am</i>	<i>pod-ōn</i>	<i>uōc-um</i>	<i>šun-ū</i>	*-ōm
dat.-abl.		<i>pad-bhyás</i>	<i>nəra-bijō</i> ³		<i>bū-bus</i> ⁴		*-bh(i)os
instr.		<i>pad-bhis</i>	<i>haō-biš</i> ⁴	<i>po-pi</i> ⁶			*-bhi(-)
loc.		<i>pat-sū</i>	<i>tanu-šū</i> ⁵	<i>po-si</i> ⁷		<i>aki-sū</i> ⁹	*-su

¹ In the phrase *kəhrpas-ca* 'and bodies'. ² In the phrase *zamas-ca* 'and lands'. ³ 'For/From men'. ⁴ 'By (ones who are) being'. ⁵ 'With bodies'. ⁶ Mycenaean; phonetically *popp̥hi*. ⁷ Synchronically labeled dative; cf. §6.11. ⁸ 'For/With cows'. ⁹ 'In the eyes'.

6.14. The animate nominative plural was *-es and the animate accusative plural was *-ns (probably from *-ms, a combination of the accusative sing. *-m and an old plural ending *-s); after consonants this became *-ns. The vocative plural was the same as the nominative, but probably with accent retraction.

6.15. The neuter nominative-accusative plural was *-h₂, originally a collective and not a true plural (see §§6.68ff.): Hitt. *āššū* 'goods' (< *h₁es-u-h₂), Ved. *paśumānt-i* 'having cattle' (-i from vocalized *-h₂), and Homeric Gk. *gēne-a* 'kinds'. (The similar-looking neuter plurals in forms like Lat. *gener-a* 'kinds' and OCS *sloves-a* 'words' actually go back to *-ā < *-eh₂ and have been taken over from the thematic declension; see §6.54 below.)

Several archaic languages also have a neuter collective plural with a lengthened grade before a resonant, such as Hitt. *widār* 'waters' and Old Av. *aiiārē* 'days'. Probably these originally ended in the collective *-h₂ too, which at an early date assimilated to the preceding resonant: *ud-or-h₂ > *ud-ōr 'waters' (> Gk. *húdōr* 'water', reinterpreted as a singular, and Hitt. *widār* with different ablaut).

6.16. The traditional reconstruction of the genitive plural ending is *-ōm. But evidence from Vedic Sanskrit and Old Avestan poetry suggests that the PIE ending was probably *-oHom, which underwent contraction to *-ōm in the later histories of the branches. (We will discuss the Sanskrit and Avestan scansion in §§10.36 and 11.21.) Some scholars have argued that there was also a short variant *-om on the basis of forms in Celtic and Slavic, but this could have arisen from *-ōm by later shortening.

6.17. The locative plural is straightforwardly reconstructible as *-su. Greek has -si, but this is probably an innovation. More problematic are the instrumental, dative, and ablative plurals. Most of the branches point to endings beginning with *-bh-, e.g. Ved. instr. pl. *pad-bhis* 'with the feet', dat. and abl. pl. *pad-bhyás* 'for the feet', Mycenaean Gk. instr. pl. *po-pi* (which spells *popp̥hi* < *pod-bhi) 'with the feet', and Lat. dat. and abl. pl. *bū-bus* 'for/with cows'. But the two northern European branches Germanic and Balto-Slavic have endings beginning with -m-, e.g. inscriptional West Germanic dat. pl. *Afli-ms* 'for the Afli (goddesses)', OCS instr. pl. *synŭ-mi* (Lith. *sūnu-mis*) 'with sons', and OCS dat. pl. *synŭ-mŭ* (Old Lith. *sūnŭ-mus*) 'for sons'. The m-endings may be an innovation of a late north-European IE dialect area.

Further confusing this picture is the fact that both **-bh-* and **-m-*elements also appear in a few cases that are *singular* and not plural in number, as Ved. *ūti-bhis* 'with help', Gk. *ī-phi* 'with might', Arm. *srti-w* 'with the heart', and OCS *synŭ-mī* 'with a son'. Additionally, they appear in *duals* in Indo-Iranian and Balto-Slavic: Ved. instr. du. *pad-bhyām* 'with the two feet', OCS dat. du. *synŭ-ma* 'for the two sons', Lith. instr. du. *nakti-m̃* 'with two nights'. Finally, neither the **-bh-* nor the **-m-*endings are found in the oldest branch, Anatolian, which has no special endings for these cases.

All this taken together suggests that the **-bh-* and **-m-*endings developed late, probably after Anatolian split off from the family, and may have originally been postpositions or adverbs ultimately related to Eng. *by* and Germ. *mit* 'with'. (It is cross-linguistically common for postpositions to develop into case-endings.)

OTHER CASE-LIKE ELEMENTS

6.18. A number of suffixes are found scattered throughout the daughter languages that may or may not be descended from true case-endings in PIE. One example is **-tos*, found in *ablative* function in both singular and plural: Ved. *pat-tās* 'from the foot', pl. *patsu-tās* 'from the feet' (interestingly, added to the locative plural), and Lat. *caeli-tus* 'from heaven'. Another is a locative ending **-dhi*, reflected by Gk. *-thi* and Arm. *-j̃* (e.g. Gk. *ouranó-thi* 'in heaven', Arm. *telowo-j̃* 'in the place'). Certain others are marginally attested and typically show up in adverbs, such as **-dhe* in Ved. *kú-ha* 'where', Av. *ku-dā* 'where', Lat. *un-de* 'whence', and OCS *kŭ-de* 'where'.

Athematic nominal stem formation

THE ACCENT-ABLAUT CLASSES

6.19. All athematic nouns consisted of three parts: root, suffix, and ending. (In one group, called root nouns, the suffix was not overt; see further below.) As already alluded to, these three morphemes could each show up in different ablaut grades depending principally on the position of the accent, which could fall on any of the three. According to the standard theory, in any given case-form of an athematic noun the unstressed morphemes appeared in the zero-grade, while the stressed morphemes were in a grade "stronger" than zero-grade – that is, one with a vowel, generally *e*, but also *o*.

6.20. Four distinct classes of athematic nouns are recognized by most scholars, each characterized by a particular patterned distribution of accent and ablaut grades. They are termed *acrostatic*, *proterokinetic*, *amphikinetic* (or *holokinetic*), and *hysterokinetic* (equivalent terms for the last three are *proterodynamic*, *amphidynamic*, and *hysterodynamic*). In all of these but *acrostatic* (which has fixed accent on the root), the accent shifts rightward in the weak cases from its position in the strong cases. Some scholars recognize two more types, *mesostatic* and *teleutostatic*, for nouns that have fixed accent on the suffix and ending, respectively; but it is not certain that these existed in the proto-language, and they will be left out of our discussion.

The somewhat simplified pictorial overview below will make the basic schemes of the four classes clear. The shaded areas indicate the location of the accent in Root or Suffix or Ending; the unshaded areas indicate no accent and usually zero-grade

(with a couple of exceptions to be treated later). An example is given beneath each class, with the nominative exemplifying the strong cases and the genitive, the weak.

	acrostatic			proterokinetic			hysterokinetic			amphikinetic		
	R	S	E	R	S	E	R	S	E	R	S	E
strong												
weak												
nom.	* <i>nók^w</i> -t-s 'night'			* <i>mén</i> -ti-s 'thought'			* <i>ph₂</i> -tér-s 'father'			* <i>h₂éus</i> -ös 'dawn'		
gen.	* <i>nek^w</i> -t-s			* <i>mñ</i> -téi-s			* <i>ph₂</i> -tr-és			* <i>h₂ús</i> -s-és		

6.21. In *acrostatic* nouns, the root is accented throughout the paradigm, but with ablaut distinction between the strong and weak cases. The most common type had *o*-grade of the root in the strong cases and *e*-grade in the weak cases. The word for 'night' above is an example; the *o*-grade nominative is preserved e.g. in Lat. *nox*, and the genitive in Hitt. *nekuz* (phonetically *nek^wts*) 'of eventide'. One well-known group of acrostatic nouns is a set of widely attested *u*-stem (§6.42) neuters such as **dóru* 'wood', **h₂óju* 'life-force', and **ǵónu* 'knee', reflected e.g. in Gk. *dóru* 'wood, spear', Ved. *áyu* 'life, life-force', and Gk. *gónu* 'knee'. Somewhat different was the word for 'liver', a neuter having lengthened grade in the nominative, **ǵékw*-r (> Av. *yākarə*, Gk. *hēpar*, whence English words like *hepatitis*).

6.22. A large number of acrostatic nouns had no overt derivational suffix ("zero suffix" in theoretical linguistics, which is not to say a non-existent suffix, but rather one without phonetic body; it is assumed that underlyingly some suffix did exist that had the function of deriving such a noun from a root, but that did not appear as a separate entity on the surface). As these nouns appear with only root and ending, they are termed *root nouns*. They include many core vocabulary items that are surely ancient, such as **h₂ek^w*- 'eye', **ped*- 'foot', and **dem*- 'house'. To take the latter as an example of the type, the nomin. sing. was **dóm*-s (> **dōm* by Szemerényi's Law [§3.38]), genit. sing. **dém*-s. The nominative singular became Arm. *toun* 'house' and probably Gk. *dō* 'house', while the genitive is found in the famous phrase **dems potis* 'master of the house, lord, master' (Ved. *dām-patis*, Old Av. *dāng paiti*-, and Gk. *des-pótēs*, the source of Eng. *despot*). We will meet other types of root nouns below.

6.23. In *proterokinetic* nouns, the root is in the full grade and accented in the strong cases, and both accent and full grade shift to the suffix in the weak cases. Most *t*- and *u*-stems (§6.42) in Sanskrit appear to have been proterokinetic, such as Ved. nomin. *matís* 'thought', accus. *matím*, genit. *matés*, from PIE **mén*-ti-s, **mén*-ti-m, **mñ*-téi-s (Vedic has generalized the zero-grade of the root throughout the paradigm; see §6.42). Neuter proterokinetic nouns behaved the same way, as in the word for 'fire', nomin.-accus. **péh₂*-uṛ, genit. **ph₂*-uén-s (see §6.32 below).

6.24. In roots ending in a resonant or resonant plus consonant, proterokinetic inflection spread at the expense of acrostatic inflection already in the proto-language. Thus alongside original acrostatic weak forms like genit. sing. **déru*-s and **dém*-s, new proterokinetic forms **dr-éu*-s and **d(ṛṇ)m-és* were created (seen e.g. in Ved. *drós* 'of the wood' and Arm. *tan* 'of the house'). Such remodelings led to a new class

of proterokinetics that descriptively had o-grade of the root in the strong cases and zero-grade in the weak. Among animate nouns of this type, the best attested is the word for 'dog', nomin. **kṛó(n)* accus. **kṛón-ŋ* genit. **kun-és*: Ved. *śvā śvān-am śún-as*, Gk. *kuón kun-á kun-ós* (with the accusative stem remade).

6.25. In hysterokinetic nouns, the suffix is accented in the strong cases, the ending in the weak. The noun for 'father' is an example: nomin. sing. **ph₂-tér-s* (the probable original form, later **ph₂-tér*; recall §6.6), accus. sing. **ph₂-tér-ŋ*, genit. sing. **ph₂-tr-és*, reflected almost exactly in Gk. *patér patéra patrós*. Another example is the noun for 'young bull', nomin. sing. **h₂uk"-s-én*, accus. **h₂uk"-s-én-ŋ*, genit. **h₂uk"-s-n-és*, reflected almost exactly in Vedic Skt. *ukṣā, ukṣānam, ukṣnás*. The English descendant is ox, whose plural *ox-en* still preserves the old *n*-stem suffix.

6.26. Root nouns with mobile accent (on the root in the strong cases, on the endings in the weak cases) are often considered as belonging to this type, though the fit is not exact: they have full grade or lengthened grade of the root in the strong cases and full grade of the ending in the weak cases. Some researchers decline to assign such nouns to any of the named accent-ablaut classes because of the lack of an overt suffix. Examples are at any rate common; they include notably the word for 'sky' or 'sky-god', nomin. sing. **diéu-s* accus. **diéu-ŋ* (which became **diém* by Stang's Law, §3.38) genit. **diu-és*, continued exactly in Ved. *dyáuṣ dyām div-ás*, Gk. *Zeús Zēn Di(w)-ós*. Hysterokinetic root nouns often appear as the second members of compounds. Such root nouns were productively formed from verbal roots to indicate either the agent or the recipient (undergoer) of an action. Examples include Ved. *rtu-íj-* 'sacrificing correctly' (root **jaǵ-* 'worship, sacrifice'), Gk. *sú-zug-* 'yoked together' (root **jeug-* 'yoke'), Lat. *con-iug-* 'spouse' (< **the one yoked with*), also from **jeug-*), and OIr. *druí*, genitive *druad* 'druid' (< **dru-uid-* 'seer/knower of the oak'; root **ueid-* 'see'). As some of these examples show, these forms often had adjectival function, and the zero-grade of the root was often generalized throughout the paradigm in the daughter languages. The hysterokinetic inflection is preserved better in Vedic, as in *vṛtra-há* 'slayer of Vṛtra', genit. *vṛtra-gbh-ás* (< PIE *-g^hhén-s *-g^hhn-és).

6.27. In amphikinetic (or holokinetic) nouns, the root is accented in the strong cases, the ending in the weak, and the suffix is typically in the lengthened o-grade (rather than the expected zero-grade) in the nominative singular and ordinary o-grade in the accusative singular. The word for 'dawn' belongs here, nomin. sing. **h₂éus-ós* (Aeolic Gk. *aúōs*, Lat. *aurōr-a*), accus. sing. **h₂éus-os-ŋ* (Ved. *uśásam*, remade with zero-grade of the root), genit. sing. **h₂us-s-és* (Ved. *uśás*). Another famous example is the word for 'earth', nomin. **dhéǵhóm* (Hitt. *tēkan*), accus. **dhéǵhom-m* (> **dhéǵhom* by Stang's Law, §3.38, to become ultimately Ved. *kṣām*), genit. **dhǵhm-és* (Hitt. *taknāš*, Ved. *ṛmās*). Yet another famous amphikinetic noun is the word for 'path', which will be discussed in §11.23.

6.28. Individual schools of thought differ on the specifics of these classes, as well as on their naturalness (cross-linguistic parallels are unfortunately lacking). The account here is based on the theory developed by German and Austrian researchers in the 1960s and 1970s (principally Jochem Schindler, Heiner Eichner, and Helmut Rix) that has become in large measure accepted in the major European and American universities. Since no branch of the family preserves the original system intact (at least in the form predicted by the theory), most of it has had to be pieced together from a variety of sometimes contradictory evidence culled from different daughter languages. It is assumed that where daughter forms are at variance with this system

(such as Ved. *matís* and *uśásam* above), the discrepancy came about through the effects of analogy and paradigm leveling that changed the position of the accent and/or the distribution of full and zero-grades within a paradigm. Further research may well lead to modifications of the theory.

INTERNAL DERIVATION

6.29. One of the most interesting facts about the above system is that the accentablaut classes in PIE are thought to have stood in a derivational relationship with one another. The language could derive new nouns or adjectives simply by shifting the accent rightward and thereby changing the form's class membership. For example, one could shift the accent of an acrostatic noun rightward and create a proterokinetic or amphikinetic noun or adjective, and from proterokinetics one could derive amphikinetics or hysteroquinetics in the same way. The resulting derivatives in general meant 'possessing, associated with' what was denoted by the original noun or adjective. This process is known as *internal derivation*, and was elucidated by the late Austrian Indo-Europeanist Jochem Schindler.

To illustrate, from an acrostatic noun could be derived a proterokinetic possessive adjective:

acrostatic **krót-u-s* genit. **krét-u-s* 'insight, intelligence' (Ved. *krátus*)

→ proterokinetic **krétus* genit. **kṛt-éu-s* 'having mental force, strong' (Gk. *kratús* 'strong')

Similarly, from a proterokinetic noun could be derived an amphikinetic noun of appurtenance:

proterokinetic **bhlégh-mṇ* 'sacred formulation' (Ved. *bráhmaṇ-*)

→ amphikinetic **bhlégh-mō(n)* 'the one of the sacred formulation, priest' (Ved. *brahmān-*)

This second type of internal derivation was quite common. Another such pair is Lat. *sēmen* 'seed' and its derivative *Sēmōn-ēs* 'deities of the seed'. From proterokinetic nouns could also be generated hysteroquinetic derivatives, as in Gk. *pseūdos* 'false' and its derivative *pseudēs* 'liar'.

6.30. Nouns could also be derived from a case-form of another noun, especially a locative. A famous example is the derivation of the word for 'man, human' from the locative of the word for 'earth'. The amphikinetic noun **dhéghōm* 'earth' had an extended locative **dhéghm-en* 'on the earth'; from this locative an amphikinetic noun meaning 'the one on/of the earth, earthling = man, human' was formed, **dhéghēmō(n)* genit. **dhéghmṇés*, becoming Lat. *homō* genit. *hominis*, and OE *guma* pl. *guman* (surviving in altered form as the *-groom* in the word *bridegroom*).

ATHEMATIC DERIVATIONAL SUFFIXES

6.31. The most prominent athematic derivational suffixes ended in *n*, *r*, *s*, and *t*, forming nouns that were called *n*-stems, *r*-stems, etc. An archaic group of nouns,

called **heteroclitic stems** (or **heteroclitcs**), had suffixes with different final consonants in the strong and weak cases. By far the best-known variety of heteroclitcs is the class of *r/n*-stems, neuter nouns ending in **-r* in the nominative/accusative singular and collective (§6.68) that is replaced by a stem in **-n-* in the other cases. A famous example is **uód-r* 'water', stem **uéd-n-*, preserved almost exactly in Hitt. *watar* genit. *witenaš*; note also that within Germanic we have both Eng. *water* (with *-r*) and ON *vatr* (with *-n*). Two other examples are Hitt. *ēšhar* 'blood' genit. *išhanaš*, and Lat. *femur* genit. *femin-is* 'thigh'. The class is widely productive in Anatolian, but moribund in the other branches. In Greek, the *-n-* was remade into *-at-*, yielding words like *húdōr* 'water' with stem *húdat-*.

6.32. A number of complex (compound) suffixes belong to this category. The richest representation of these is again provided by Anatolian, which reflects stems in **-mer/n-*, **-ser/n-*, **-ter/n-*, and **-uer/n-*. The last of these is found for instance in the word **peh₂-u₁* 'fire', continued most faithfully in Hitt. *pahhur* genit. *pahhuenaš*; the alternation is preserved indirectly in Arm. *hour* 'fire' alongside its derivative *hn-oc* 'oven' (< **hown-oc*) with *-n-*, and in Eng. *fire* alongside Goth. *fon* 'fire'.

6.33. There may also have been an *l/n*-stem, the word for 'sun': nomin. **seh₂-uól* (in e.g. Lat. *sōl*, Goth. *sauil*), genit. **sh₂-uén-s* (OAv. *x^oəng*) or **sh₂-un-s* (Eng. *sun* is ultimately from a similar stem). But given that *sōl* etc. is masculine while *x^oəng* is neuter, these may rather be from separate paradigms.

6.34. *n*-stems. Various types of animate and neuter suffixes ending in **-n* have been reconstructed, especially the compound suffixes **-men-*, **-sen-*, **-ien-*, and **-uen-*, all of which frequently show up as the basis for infinitival endings (see §5.54). The most familiar is the neuter abstract noun suffix **-m₁*, genit. sing. **-men-s*, referring to the act or result of the action of the verb. This yields for example the Vedic and Greek abstract nouns in *-ma* (e.g. Ved. *kár-ma* 'deed', lit. 'thing done', Gk. *thé-ma* 'theme', lit. 'thing placed') and the Latin abstracts in *-men* (e.g. Lat. *car-men* 'song' < **kan-m₁* 'thing sung'). Some more archaic nouns consisted simply of an *n* preceded or followed by an ablauting vowel, such as **g^wreh₂u-ō(n)* 'pressing stone, millstone' (from **g^wh₂u-* 'heavy', as in Ved. *gurú-*), becoming Ved. *grāvān-*, OIr. *bráun* (genit. *broon*), and Eng. *quern* 'hand mill'.

6.35. The so-called "Hoffmann" suffix (named after the late Indologist and Indo-Europeanist Karl Hoffmann) has been recently added to the dossier. Forming adjectives that indicated possession, it had the shape **-Hon-* alternating with zero-grade **-Hn-*. The PIE adjective for 'young' is a well-known example, made from the zero-grade of **h₂oju* 'life-force': nomin. **h₂iu-Hō(n)* (Ved. *yúvā*), accus. **h₂iu-Hon-₁* (Ved. *yúvānam*), genit. **h₂iu-Hn-es* (Ved. *yūn-as*). This adjective thus originally meant 'having life-force, having vitality'.

6.36. A very widespread neuter *n*-stem noun is the word for 'name', but its precise reconstruction is disputed; two commonly adduced contenders are **h₂neh₂m₁* and **h₁nom₁*. The cognates include Hitt. *lāman* (dissimilated from earlier **nāman*), Ved. *nāma*, Greek *ónoma* (Laconian Gk. *énuma*, only in personal names and preserving an original *e-* that changed to *o-* elsewhere in Greek), Lat. *nōmen*, Goth. *namo*, Arm. *anoun*, OIr. *ainm(m)*, Toch. A *ñom*, and OCS *imę*.

6.37. *r*-stems. Many stems ended in an **-r* that were not heteroclitic. The most common were the agent-noun suffixes **-ter-* and **-tor-*. The first of these formed so-called non-event agent nouns, as in **dh₃-tér* 'giver, one whose function or role is to give (but who may never have actually done so)' (Ved. *dātā*, Gk. *dotér*), while

the second formed event agent nouns (**déh₃-tōr* 'giver, one who has given': Ved. *dātā*, Gk. *dōtōr*). Besides their different semantics, the two behaved differently syntactically, too: in Indo-Iranian the first type took a genitive complement (Ved. *dātā vāsūnām* [genit. pl.] 'giver of goods'), whereas the *-*tor*- nouns took an accusative (*dātā vāsūni*).

Another suffix *-*ter*- is found in kinship terms: **ph₂ter*- 'father', **mater*- 'mother', **bhrāter*- 'brother', **dhugh₂ter*- 'daughter', and **jenh₂ter*- 'husband's brother's wife'. Attempts to analyze these words as agent nouns have been unconvincing.

An *r*-stem with simpler stem is the word for 'hand', **ǵh₃s-ōr* stem **ǵhes-r*-, seen in Hittite *kiššaraš* (phonetically *kisars*; the *-*s* was added later) stem *kiš(š)r*-, Gk. *kheir*-, and Arm. *jejn* (< **ǵhesr-n*-).

6.38. *s*-stems. Neuters ending in *-*os* in the strong cases and having a stem *-*es*- in the weak cases, but with full grade of the root throughout (and therefore not falling readily into any of the accent-ablaut classes), are a common type of verbal abstract noun and were formed from the full grade of verbal roots. Thus from **ǵenh₁-* 'be born' was formed **ǵénh₁-os* 'birth, thing born, race, kind', stem **ǵenh₁-es*-. Ved. *jānas*- 'race, clan', Gk. *génos* 'race, kind' (genitive *géne-os* < pre-Greek **gen-es-os*), Lat. *genus* 'type', and OIr. *cin* 'birth'.

A simple suffix *-*s*- is found in some archaic nouns, such as **kreyh₂-s* 'gore, bloody flesh' (Ved. *krāviš*-, Gk. *kréas*; the same root that yields English *raw*).

6.39. Animate *s*-stems having possessive semantics were formed from neuter *s*-stems by internal derivation (§6.29). Thus the neuter noun for 'grain', **kerh₁-os* **kerh₁-es*-, preserved in German *Hirse* 'millet', could form a noun or adjective **kerh₁-és* '(the one) associated with or possessing grain', whence the name of the Latin agricultural goddess *Cerēs*. Similarly, the neuter Greek noun *klé(w)os* 'fame' had a derivative *-klé(w)ēs* (contracted to *-klēs*) 'famed', common in names like *Sopho-klēs* 'famed for wisdom'.

A special *s*-stem was the suffix *-*uos*- used to form the perfect participle (§5.56).

6.40. *t*-stems. Athematic suffixes in *-*t*- are best known for forming feminine abstract nouns. The simple suffix is found for example in PIE **dekṃ-t*- 'group of ten, decad' (from **dekṃ* 'ten'): Ved. *daśāt*- and OCS (pl.) *desęte*. It is also found in the compound suffixes *-*tāt*- (probably from earlier *-*teh₂-t*-) and *-*tūt*- (*-*tuh₂-t*-), likewise used to form feminine abstracts; these are most familiar from Latin: *libertāt*- 'freedom', *senec-tūt*- 'old age'. An unrelated suffix of some interest is *-*it*-, seen in PIE **melit*- 'honey' and Hitt. *šēppit*- 'kind of grain'; perhaps it was a "foodstuff" suffix.

An important subset of the *t*-stems is comprised by the *nt*-stems, whose most conspicuous members are the participles formed to present and aorist tense-stems (see §5.56). Also worth mentioning is the possessive suffix *-*yent*-, as in Hitt. *ēšhar-want*- 'bloody' (*ēšhar* 'blood'), Ved. *bhāga-vant*- 'wealthy' (*bhāgas* 'wealth'), and Gk. *niphó-(w)ent*- 'snowy' (*niph*- 'snow').

6.41. Reduplicated athematic nouns. A variety of reduplicated nouns are known from all stem-classes. One reconstructible for PIE is the word for 'beaver', **bhebhru*- (a derivative of **bhreun*- 'brown'): OHG *bibar*, Lith. *bēbrus*, Ved. *babhrú*- 'brown'.

6.42. *i*- and *u*-stems. Two classes of athematic nouns that behaved in parallel were those whose suffix was *-*i*- or *-*u*- in the zero-grade and *-*ei*- *-*eu*- in the full grade. The most common type continues an earlier proterokinetic paradigm, with zero-grade *-*i-s* *-*i-m* and *-*u-s* *-*u-m* in the strong cases, and full grade in the

weak cases (genit. **-ei-s* **-eu-s*). By the time of the daughter languages, however, these nouns do not ablaut in their root syllable. Thus Ved. *matī-* 'thought' is believed to continue a proterokinetic nomin. **mén-ti-s* genit. **mṛ-tēi-s*, but the zero-grade **mṛ-* (> Ved. *ma-*) has been generalized. (The same happened in Greek and other branches.) This is an example of the verbal abstract nouns in **-ti-*, one of the most common groups of *i*-stems; also common were abstract nouns in **-tu-*. Both of these types often appear in infinitives in the daughter languages (see §5.54).

Many adjectives made from roots having inherently adjectival meaning are simple *u*-stems, such as Lith. *dubūs* 'deep' and Ved. *gurú-* 'heavy'. On the *u*-stem proterokinetic nouns of the type **dóru* 'spear', see §6.21.

Thematic Nouns

6.43. Thematic nouns, like thematic verbs, are characterized by the presence of an ablauting stem **-e/o-* to which the case-endings are added. In practice, the stem is almost always **-o-*, and so these nouns are commonly referred to as *o*-stems. The case-endings are mostly the same as the athematic ones, but there are several notable differences, as we will see shortly. In many of the daughter languages, certain thematic case-endings were influenced or replaced by the corresponding pronominal case-endings (see §7.8); some common instances of this will be pointed out.

6.44. There are no certain accent-ablaut distinctions within thematic paradigms except in the vocative, where the accent was retracted. (But see also §6.66.) Most animate *o*-stems had masculine gender, but a number of them, including some kinship terms (e.g. **snusos* 'daughter-in-law') and a variety of tree names (e.g. **bhāḡos* 'oak, beech' and **kṛnos* 'cornel cherry'), were feminine.

Case-endings of thematic nouns

6.45. Below is a comparative table of the athematic and most archaic thematic endings to illustrate the similarities and differences between the two classes:

	Singular		Dual		Plural	
	athem.	them.	athem.	them.	athem.	them.
nom. anim.	<i>*-s</i>	<i>*-os</i>	<i>*-h₁</i>	<i>*-ō < *-o-h₁</i>	<i>*-es</i>	<i>*-ōs < *-o-es</i>
voc.	<i>*-ø</i>	<i>*-e</i>	<i>*-h₁</i>	<i>*-ō < *-o-h₁</i>	<i>*-es</i>	<i>*-ōs < *-o-es</i>
acc. anim.	<i>*-m</i>	<i>*-om</i>	<i>*-h₁</i>	<i>*-ō < *-o-h₁</i>	<i>*-ns</i>	<i>*-ons</i>
n.-a. neut.	<i>*-ø</i>	<i>*-om</i>	<i>*-ih₁</i>	<i>*-o-ih₁</i>	<i>*-h₂</i>	<i>*-ā < *-e-h₂</i>
gen.	<i>*-s</i>	<i>*-os</i> ?			<i>*-ōm</i>	<i>*-ōm</i>
abl.	<i>*-s</i>	<i>*-ōt < *-o-(h₂)at</i>			<i>*-bh-</i>	<i>*-o(i)bh-</i>
dat.	<i>*-ei</i>	<i>*-ōi < *-o-ei</i>			<i>*-bh-</i>	<i>*-o(i)bh-</i>
instr.	<i>*-h₁</i>	<i>*-ō < *-o-h₁</i>			<i>*-bh-</i>	<i>*-o(i)bh-</i>
loc.	<i>*-i</i>	<i>*-oi</i>			<i>*-su</i>	<i>*-oisu</i>

SINGULAR

6.46. The following chart shows some of the comparative evidence for the singular declension of *o*-stems: the animates Hitt. *arunaš* 'sea', Ved. *vīrás* 'man', Av. *yasna-*

'sacrifice', Gk. *theós* 'god', Lat. *servus* 'slave', Goth. *dags* 'day', OCS *rabŭ* 'slave', and Lith. *výras* 'man'. The neuters and other special forms are glossed in the notes.

	Hitt.	Ved.	Av.	Gk.	Lat.	Goth.	OCS	Lith.
nom. anim.	<i>arun-aš</i>	<i>vīr-ás</i>	<i>yasnō</i>	<i>the-ós</i>	<i>servu-os</i> ⁹	<i>dag-s</i>	<i>rab-ŭ</i>	<i>výr-as</i>
voc. anim.	<i>negn-a</i> ¹	<i>vīr-a</i>		<i>ádelph-e</i> ⁶	<i>servu-e</i>		<i>rab-e</i>	<i>výr-e</i>
acc. anim.	<i>arun-an</i>	<i>vīr-am</i>	<i>yasn-am</i>	<i>the-ón</i>	<i>servu-om</i> ⁹	<i>dag</i>		<i>výr-a,</i> <i>denu-an</i> ¹⁶
n.-a. neut.	<i>pēd-an</i> ²	<i>pad-ám</i> ³	<i>xšaθr-am</i> ⁵	<i>pēd-on</i> ⁷	<i>pers-om</i> ¹⁰			<i>gīvu-an</i> ¹⁷
gen.	<i>arun-aš</i>	<i>vīr-asya</i>	<i>yasn-ahe</i>	<i>the-oío</i> ⁸	<i>servu-i,</i> <i>Popli-osio</i> ¹¹			
					<i>Gnau-ōd</i> ¹²		<i>rab-a</i> ¹³	<i>výr-o</i> ¹⁵
abl.		<i>vīr-āt</i>	<i>yasn-āi</i>		<i>Herakl-ōi</i> ¹³	<i>dag-a</i>	<i>rab-u</i>	<i>výr-u</i>
dat.		<i>vīr-āya</i>	<i>yasn-āi</i>	<i>the-ōi</i>				<i>výr-u</i>
instr.		<i>yajñ-ā</i> ⁴	<i>yasn-a</i>					
loc.	<i>arun-i</i>	<i>vīr-é</i>	<i>yasn-e</i>		<i>dom-i</i> ¹⁴		<i>rab-ē</i>	<i>nam-iē</i> ¹⁴

¹ 'Brother'. ² 'Place'. ³ 'Track'. ⁴ 'With a sacrifice'. ⁵ 'Rule'. ⁶ 'Brother'. ⁷ 'Ground'. ⁸ Homeric; *theōi* in Classical Gk. ⁹ Archaic Lat.; *servus* and *servum* in Classical Lat. ¹⁰ Umbr., 'ground'. ¹¹ Archaic Lat., 'of Publius'. ¹² Archaic Lat., 'from Gnaeus'. ¹³ Osc., 'for Hercules'. ¹⁴ 'At home'. ¹⁵ Synchronically genitive. ¹⁶ OPruss., 'god'. ¹⁷ OPruss., 'alive'.

6.47. There is complete agreement on the reconstruction of the animate nominative, vocative, accusative, and neuter nominative-accusative singular. Retraction of the accent in the vocative is evidenced by such forms as Ved. *vīr-a* (vs. nomin. *vīr-ás*) and Gk. *ádelph-e* (vs. nomin. *ádelphós*) above. Also uncontroversial are the reconstructions of the dative, instrumental, and locative singular, traditionally reconstructed as **-ōi*, **-ō*, and **-oi*. The first two are likely contracted from **-o-ei* and **-o-h*. An ablaut variant **-ei* in the locative is found especially in Italic, e.g. Oscan *comenei* 'in the assembly'.

6.48. The picture for the genitive singular is not quite as clear. The oldest ending was perhaps **-os*, identical to the nominative, if Anatolian is any indication (e.g. Hitt. nomin. *arun-aš* 'sea', genit. *arun-aš*); and this ending may have infiltrated the athematic declension in several daughter languages, including Greek (see §6.8). But much more widely represented is **-osjo*, represented by Ved. *vīr-asya*, Av. *yasn-ahe*, Homeric Gk. *the-oío*, and Archaic Lat. *Popli-osio* above, as well as by Arm. *get-oy* 'of the river'. It has been proposed that **-osjo* is a combination of the earlier **-os* and the stem of the relative pronoun **iō-* (§7.10); in many languages, relative pronouns are used to indicate possession. A similar suffix **-oso* is reflected by OE (Northumbrian) *moncynn-æs* 'of mankind' (> Mod. Eng. *-s*) and probably OPruss. *deiū-as* 'of god'; it is related to the pronominal genitive **-eso* (see §7.8). A completely different genitive ending **-i* is found in Italic (Latin *uir-i*), the closely allied Venetic (*Vineti* 'of Vinetos'), Celtic (Ogam Ir. *maq(q)i* 'of the son'), and probably Messapic (*Dazimaihi* 'of Dazimas'; *-aihi* spells *-ai* < **-o-i*).

6.49. Unlike the athematic ablative, the thematic ablative singular is distinct from the genitive. It was formed by suffixing an element variously reconstructed as **-ad*, **-at*, or **-h₂et* (> **-h₂at*) to the thematic vowel (the final consonant is uncertain due to the word-final voicing neutralization discussed in §3.40). This yielded a sequence **-o-(h₂)ad/t* that became *-āt* in Vedic and *-āt* in Old Avestan, and elsewhere contracted to **-ōt* (e.g. Archaic Lat. *Gnau-ōd* 'by Gnaeus') or **-āt* in Balto-Slavic, where it became a genitive (OCS *grad-a* 'of/from the city', Lith. *výr-o* 'of/from the man'). This suffix may be related to the Slavic preposition **otŭ* 'from' (> Russ. *ot*).

DUAL

6.50. The nominative, vocative, and accusative animate dual is traditionally reconstructed as **-o*, which is now thought to be from **-o-h₁* on the assumption that the ending contains the same dual suffix **-h₁* seen in the athematic declension. Examples of the thematic dual include Gk. *anthrōp-ō* 'two men', Ved. *dev-ā* 'two gods', OCS *grad-a* 'two cities', and Lith. *vyr-u* 'two men'. The thematic neuter nom.-voc.-acc. dual ended in **-o-ih₁*, as in Ved. *yug-é* and OCS *iz-ě* 'two yokes'. The other cases in the dual cannot be reconstructed with any certainty.

PLURAL

6.51. The following chart shows some of the comparative evidence for the plural of o-stems, using the same nouns as above except for Ved. *devás* 'god' and Av. *daēnuha-* 'demon'. Other forms are glossed in the notes.

	Hitt.	Ved.	Av.	Gk.	Lat.	Goth.	OCS	Lith.
nom. anim.		<i>dev-ās</i>	<i>daēnu-āṇhō</i>	<i>the-oi</i>	<i>seru-i</i>	<i>dag-os</i>	<i>rab-i</i>	<i>vyr-ai</i>
acc. anim.	<i>arun-us</i>	<i>dev-ān</i>	<i>mašu-šug</i> ³	<i>the-oīs</i> , <i>odel-ons</i> ⁶	<i>seru-ōs</i>	<i>dag-ans</i>	<i>rab-y</i>	<i>vyr-us</i> , <i>deinu-ans</i> ¹⁰
neut. n.-a.	<i>pēd-a</i> ¹	<i>yug-ā</i> ²		<i>zug-ā</i> ³	<i>locā</i> ¹			<i>slay-o</i> ¹¹
gen.	<i>arun-an</i>	<i>dev-ānām</i>	<i>daēnu-anam</i>	<i>the-ōn</i>	<i>seru-om</i> ⁷		<i>rab-ū</i>	<i>vyr-ū</i>
dat.		<i>dev-ēbhryas</i>	<i>daēnu-aēbriō</i>		<i>Andetic-obos</i> ⁸	<i>dag-am</i>	<i>rab-omū</i>	<i>vyr-ams</i>
instr.		<i>dev-ās</i> , <i>dev-ēbhīs</i>	<i>yasnās</i> ⁹	<i>the-oīs(ī)</i>	<i>seruīs</i> , <i>sokius</i> ⁹			<i>vyr-ais</i>
loc.		<i>dev-ēṣu</i>	<i>ašp-aēṣu</i> ⁵				<i>rab-ēxū</i>	

¹ 'Places'. ² 'Yokes'. ³ OAv., 'mortals'. ⁴ 'With sacrifices'. ⁵ 'Horses'. ⁶ Cretan, 'obols'. ⁷ Archaic Lat. ⁸ Venetic, 'for the sons of Andetios' (?). ⁹ Archaic Lat. (inscriptional), 'for the friends'. ¹⁰ OPruss., 'gods'. ¹¹ OPruss., 'sled' (pl. in form but sing. in meaning).

6.52. The animate nominative and vocative plural was **-ōs* (contracted from **-o-es*), as in Ved. *dev-ās* 'gods', OIr. (*á*) *fir-u* 'men' (surviving as a vocative only), Goth. *dag-os* 'days', and Osc. *Núvian-ús* 'men from Nola' (the boldfacing is customary for transliterating words written in the native Oscan alphabet). To judge by the Ved. vocative *dēvās* 'gods!', the vocative plural had accent retraction like the singular. A doubled ending **-ōsos* is found in Indo-Iranian and Germanic: Ved. *dev-āsas* 'gods', Av. *daēnu-āṇhō* 'demons', and probably East Frisian *fisk-ar* 'fishes'.

6.53. In at least five branches the pronominal nominative plural **-oi* (§7.8) replaced the original nominal ending. This probably came about in phrases consisting of a pronoun and noun, where the noun simply copied the pronominal ending: so a phrase **toi ekuōs* 'those horses, the horses' became **toi ekuoi*. The ending is seen in Gk. *the-oi*, Lat. *seru-i*, OCS *rab-i*, and Lith. *vyr-ai* above, as well as Toch. B *yakwi* 'horses' (< **ekyoi*) and OIr. *fir* 'men' (< **uir-oi*).

6.54. The neuter nominative-accusative plural was **-eh₂*, originally a collective (§§6.68ff.), as in Ved. *yug-ā* 'yokes' and Osc. *comono* 'assemblies' (with *-o* < **-ā*).

6.55. The other thematic plural endings resemble the athematic endings, and generally consist of them added to the thematic vowel. The animate accusative plural was **-ons* (= **-o-ns*), as seen in OAv. *mašu-šug*, Gk. (Cretan) *odel-ons*, Goth.

dag-ans, and OPruss. *deiū-ans* above. A long version **-ōns*, which may or may not be of PIE date, is reflected for example in Ved. *dev-ān* (appearing as *dev-āns* before certain consonants) and Umbr. *uer-of* 'gate(s)' (-f in Umbrian is the regular outcome of **-ns*!).

6.56. The thematic genitive plural ended like the athematic genitive in **-ōm*, as seen in Gk. *the-ōn*, Lith. *výr-ų*, and maybe OPruss. *grec-on* 'of sins'. This was apparently shortened to **-om* in Celtic, perhaps Archaic Lat. *uir-um* 'of men', and OCS *rab-ŭ*.

6.57. As in the athematic declension, the thematic dative, instrumental, and ablative plural were generally formed with endings in **-bh-* or **-m-*, the latter only in Germanic and Balto-Slavic. Sometimes, as in Ved. instr. pl. *dev-ébbis* and Av. dat. pl. *daēuu-aēbiiō*, the ending was added to a stem in **-oi-* rather than the bare thematic stem. This feature resurfaces elsewhere: in the one reconstructible instrumental ending that has no parallel in the athematic declension, **-ōis* (reflected e.g. in Ved. *dev-āis*, Lith. *výr-ais*, and Gk. *the-oīs* [dative]), and in the thematic locative plural ending **-oi-su* (Ved. *dev-ēsu*, Av. *asp-aēšu*, Homeric Gk. *the-oīsi* [dative], and OCS *rab-ěxŭ* with *-x-* from earlier **-s-*).

Thematic nominal formations and suffixes

6.58. Some thematic nouns reconstructible for PIE are not derived from any known root, such as **u₁lk^wos* 'wolf', **h₂f^{ik}os* 'bear', **snusós* 'daughter-in-law', and **ag^wnos* 'lamb'. But most of the thematic nouns found in the daughter languages are derived from roots, either by adding the thematic vowel alone or by adding a suffix that itself ends in the thematic vowel. PIE had a large number of such suffixes, a few of which are given below. Some of the example forms are actually feminines ending with the suffix **-ā* (**-e-h₂*; see §§6.69–70).

TOMÓS AND TÓMOS NOUNS

6.59. These nouns, named after a Greek example of each type, were formed by adding the thematic vowel to the o-grade of the root. *Tomós* nouns had accented suffix and agential or active meaning (Gk. *tomós* means 'sharp', i.e. 'cutting', from **tem-* 'cut'), while *tómos* nouns had root accent and resultative meaning (*tómos* means 'a slice', that is, a thing that results from cutting). The agential *tomós* nouns are particularly frequent as final compound members; an example is **-g^wor(h₃)os* 'eating, devouring' (root **g^werh₃-* 'devour') in Gk. *dēmo-bóros* 'devouring the people' and Lat. *carni-uorus* 'eating meat, carnivorous'.

POSSESSIVE DERIVATIVES IN **-ó-*

6.60. Thematic nouns and adjectives indicating possession could be formed by adding the accented thematic vowel suffix **-ó-* to an athematic noun stem, or by shifting the accent rightward to the thematic vowel of an already existing thematic noun (usually with additional modifications, see directly below). For example, the noun **g^wieh₂-* 'bowstring' (> Ved. *jyā*) had zero-grade **g^wih₂-*, from which a thematic noun **g^wih₂-ó-* '(thing) having a bowstring' could be formed (> Gk. *biós* 'bow').

6.61. At other times, the ablaut grade of the base was "increased" a step (from zero to full or from full to lengthened), a very common type of formation called a *vrddhi-*

derivative, after the Sanskrit word for 'growth'. Vrddhi-derivatives have a genitival or ablational relationship to the noun from which they are formed: 'belonging to, of, (descended) from'. Thus from the thematic noun **syékuro-* 'father-in-law' (Ved. *śvāsura-*) could be formed a lengthened-grade vrddhi-derivative **syékuró-* 'pertaining to one's father-in-law' (> Skt. *śvāsura-* 'relating to one's father-in-law', OHG *swāgur* 'father-in-law's son = brother-in-law').

6.62. Sometimes a vrddhi-derivative was formed from a zero-grade and involved the insertion of the full-grade vowel in the "wrong" place. A famous example is the PIE word for 'god', **deiy-ó-s* (Ved. *devás*, Av. *daēuua-* 'demon', Archaic Lat. accus. pl. *deiuōs* 'gods'). It was formed from the zero-grade **diu-* of the noun **d̥iēus* 'sky', with the "wrong" full-grade **deiy-* instead of **d̥iēu-*. Its literal meaning was 'one of, belonging to, or inhabiting the sky'.

REDUPLICATED THEMATIC NOUNS

6.63. Just as some athematic nouns exhibit reduplication, so too do some thematic nouns. The most famous example, found throughout the family, is **k^we-k^wl-o-* 'wheel' (root **k^wel(H)-* 'to turn'): Ved. *cakráṁ*, Gk. *kúklos*, Toch. A *kukäl* 'chariot', OE *hweohl* 'wheel', and perhaps Hitt. *kugullaš* 'ring-shaped bread, donut'. Additional examples include Ved. *ré-rih-a-* 'licking repeatedly', Av. *ṛa-rem-a-* 'resting', and OCS *gla-gol-ŭ* 'word' (< **gal-gal-o-*, from **gal-* 'to cry out'). It has been suggested that such forms had intensive or emotive force; **k^we-k^wl-o-* may have been an expressive neologism for a new gadget, the wheel, referring to the iterativity of its turning. (Recall on this point also §2.58.)

NOUNS IN **-lo-*

6.64. This is widely found in diminutives in Italic, Germanic, and Balto-Slavic, such as the word for 'piglet': Lat. *porc-ulus*, OHG *farh-ali*, and Lith. *parš-ėlis* (only the Latin form has actually remained thematic). It is also found added to verbal roots to make nouns of various kinds, such as **sed-lo-* (or feminine **sed-leh₂*) 'a seat' (from **sed-* 'sit') in e.g. Gk. *hellá*, Lat. *sella*, and Goth. *sittls*.

NOUNS IN **-mo-*

6.65. This suffix was added to verbal roots to form a variety of action and event nouns. For example, the root **g^wher-* 'to heat up' formed **g^wher-mo-* or **g^whor-mo-*, appearing as a noun meaning 'heat' (Ved. *gharmá-*, OPruss. *gorme*, Alb. *zjarm* 'fire') or an adjective meaning 'warm' (Gk. *thermós*, Lat. *formus*, Eng. *warm*, Arm. *ǰerm*). Some further examples from the individual languages include Gk. *gnó-mē* 'opinion' (Greek root *gnō-* 'know' < PIE **gneh₃-*), Lat. *fā-ma* 'fame, report, reputation' (Latin root *fā-* 'speak' < PIE root **bheh₂-*), and Eng. *stream* (< **sreu-mo-*, from **sreu-* 'flow').

On the participial suffix **-mo-*, see §5.56.

NOUNS IN **-no-*

6.66. This suffix was added to verbal roots to form nouns of various sorts, such as Vedic *yajñā-* and Av. *yasna-* 'sacrifice, offering' (< **iaǵ-no-*, from **iaǵ-* 'to sacrifice'),

and Lat. *donum* 'gift' (< **deh₃-no-*, from **deh₃-* 'to give'). One well-known trio of these nouns – the words for 'sleep' (**syep-no-*), 'wagon' (**uegh-no-*), and 'price' (**ues-no-*) – is unusual in showing ablaut of the root. To use the first of these for illustration, the *e*-grade **syep-no-* is found in OE *swefn* 'sleep, dream', the *o*-grade **syop-no-* in Arm. *k'own* 'sleep', and the zero-grade **sup-no-* in Gk. *húpnos* 'sleep'.

On the passive verbal adjective suffix *-*nó-*, see §5.57.

"TOOL" NOUNS IN *-*tlo-*, *-*dhlo-*, *-*tro-*, *-*dhro-*

6.67. These suffixes formed nouns indicating an instrument for accomplishing an action; normally they are neuters. Originally there were only the suffixes *-*tlo-* and *-*tro-*; *-*dhlo-* and *-*dhro-* arose secondarily by Bartholomae's Law (§§3.35, 10.6) and were later reanalyzed as independent suffixes. Examples of each suffix include *(s)*neh₁-tlo-* 'thing to sew with' (from **sneh₁-* 'sew') in Eng. *needle*; *(s)*th₂-dhlom* 'place to stand things' (from **steh₂-* 'stand') in Lat. *stabulum* 'place for animals to stand, stable, stall'; **h₂erh₃-trom* 'a plow' (from **h₂erh₃-* 'to plow') in Gk. *árotron*, Lat. *arātrum*, and OIr. *arathar*; and **kr(e)i-dhro-* 'sieve' (from **krei-* 'to sift, strain') in Lat. *cūbrum* and OE *briddar*.

The Collective and the Feminine

6.68. As noted above, the neuter plural was formed by adding *-*h₂*. A curious fact in Anatolian, Greek, and Old Avestan is that neuter plurals took singular verb agreement, as in the Greek phrase *tà zōia trékhei* 'the animals (pl.) run (sing.)'. Additionally, in several branches (most notably Anatolian), animate nouns could form collective plurals in -*a* (< *-(e)*h₂*) alongside ordinary (distributive) plurals: Hitt. *alpaš* 'cloud', ordinary (distributive) plural *alpeš* 'clouds', collective *alpa* 'group of clouds'; Gk. *mēros* 'thigh', ordinary plural *mēroi* '(individual) thigh-pieces' (of a sacrificial animal), collective plural *mēra* '(group of) thigh-pieces'; Lat. *locus* 'place', ordinary plural *loci*, collective *loca*. These facts have led many to hypothesize that *-*h₂* was originally a collective suffix, and that PIE possessed a singular, dual, distributive plural, and collective.

6.69. The suffix *-*h₂* is at least outwardly identical to an animate suffix *-*h₂* that forms feminine nouns and abstract nouns. Its function as an abstract-noun suffix is found in all the branches including Anatolian, as in Luvian *zid-āh(-iša)* 'virility' (from *ziti-* 'man'), Lycian *pijata-* 'gift' (< **pñjo-teh₂*), and Gk. (Doric) *tom-ā* 'a cutting' (< **tom-eh₂*). The feminizing *-*h₂*, however, is found only outside Anatolian. If these two *-*h₂*-suffixes are one and the same, as most researchers believe, then the collective *-*h₂* early on acquired the additional function of marking the animate singulars of abstract nouns, and then later (after Anatolian broke off from the family) became specialized as a feminizing suffix.

It is not clear how this development is to be explained semantically. The use of abstract nouns to refer to an individual person belonging to or associated with that abstract entity has good parallels, as in Eng. *youth* 'state of being young' > 'a young individual', or Lat. *testimōnium* 'testimony' > French *témoin* 'witness'. But how such individualizing use might have been manifested in feminizing function is unclear.

The collective *-*h₂* was restricted to the nominative and accusative of collectives, but in abstract and feminine forms it is found in all cases and numbers.

The *ā*-stems

6.70. The most widespread group of feminine nouns had a stem **-eh₂-*, which contracted to **-a-*, the source of the *ā*-declensions that are characteristic of feminine nouns in so many of the daughter languages (Greek and Latin among them). This stem is either the feminizing suffix **-h₂* added to the thematic vowel, or simply the full grade of the feminizing suffix. Compare the paradigms below of Ved. *priyā* 'dear', Av. *daēna* 'religion', Gk. *theā* and Lat. *dea* 'goddess', Goth. *giba* 'gift', OCS *glava* 'head', and Lith. *dūona* 'gift'.

Singular

	Ved.	Av.	Gk.	Lat.	Goth.	OCS	Lith.
nom.	<i>priy-ā</i>	<i>daēn-a</i>	<i>the-ā</i>	<i>de-a</i>	<i>gib-a</i>	<i>glav-a</i>	<i>dūon-a</i>
acc.	<i>priy-ām</i>	<i>daēn-am</i>	<i>the-ān</i>	<i>de-am</i>	<i>gib-a</i>	<i>glav-ŏ</i>	<i>dūon-ą</i>
gen.	<i>priy-āyās</i>	<i>daēn-aiiā</i>	<i>the-ās</i>	<i>de-ae,</i> <i>de-āi</i> ¹	<i>gib-os</i>		<i>dūon-os</i>
dat.	<i>priy-āyai</i>	<i>daēn-aiiāi</i>	<i>the-āi</i>	<i>de-ae</i>	<i>gib-ai</i>	<i>glav-ě</i>	<i>dūon-ai</i>
instr.	<i>priy-ā,</i> <i>priy-āyā</i>	<i>daēn-a,</i> <i>daēn-aiia</i>				<i>glav-ojo</i>	<i>dūon-a</i>
loc.	<i>priy-āyām</i>	<i>daēn-a</i>				<i>glav-ė</i>	<i>dūon-oje</i>

Plural

nom.	<i>priy-ās</i>	<i>daēn-ā</i>	<i>the-ai</i>	<i>de-ae</i>	<i>gib-os</i>	<i>glav-y</i>	<i>dūon-os</i>
acc.	<i>priy-ās</i>	<i>daēn-ā</i>	<i>the-ās</i>	<i>de-ās</i>	<i>gib-os</i>	<i>glav-y</i>	<i>dūon-as</i>
gen.	<i>priy-ānām</i>	<i>daēn-anam</i>	<i>the-ōn,</i> <i>the-ōn</i> ²	<i>de-ārum</i>	<i>gib-o</i>	<i>glav-ŭ</i>	<i>dūon-ų</i>
dat.	<i>priy-ābhyas</i>	<i>daēn-ābiio</i>	<i>the-āis(i)</i>	<i>de-is</i>	<i>gib-om</i>	<i>glav-amŭ</i>	<i>dūon-oms</i>
instr.	<i>priy-ābhis</i>	<i>daēn-ābiš</i>	<i>the-āphī</i> ²	<i>de-ābus</i> ³		<i>glav-amu</i>	<i>dūon-omis</i>
loc.	<i>priy-āsu</i>	<i>daēn-āhu</i>				<i>glav-axŭ</i>	<i>dūon-ose</i>

¹ Archaic Lat.; note also the fixed phrase *pater famili-ās* 'head of the household'. ² Homeric.

³ Synchronically dative and ablative.

As can be seen, the **-eh₂-*declension is athematic, with the case-endings added directly to the laryngeal. There is interestingly no **-s* in the nominative singular. Some of the forms in the charts above have been remade, such as the Latin genitive sing. *de-ae* (§13.38). Note that forms like the accusative singular **-eh₂-m* do not appear as **-eh₂-m*, as might be expected from the usual syllabification rules (§3.42), but rather remain **-eh₂-m*, eventually to become **-ām*. This is an example of Stang's Law (see §3.38).

At an earlier stage there were potentially feminines formed simply by adding **-h₂* to a root, as in **g^wen-h₂* 'woman', reflected directly by Ved. *jāni-* (with *-i-* < **h₂*). (Gk. *gunē* and OCS *žena* come from later forms ending in **-eh₂-*.)

The *ā*-declension in the daughter languages is familiarly associated with thematic declensions, especially in adjectives where *o*-stem masculine and neuters form their feminines with **-a-* (e.g. Lat. *nouns* [masc.], *nouum* [neut.], *noua* [fem.]). It is also associated with feminine counterparts of *o*-stem animates (e.g. *ursa* 'she-bear' next to *ursus* 'he-bear' in Latin). Both of these functions, however, may have arisen relatively late, as they are barely found in the earliest attested stages of these languages.

The *devī-* and *vr̥kī-*feminines and other formations

6.71. Beside bare **-eh₂-* we can reconstruct a compound suffix **-ieh₂-* that is also collective and feminizing. It existed in two types, represented on the one hand by Vedic *devī* 'goddess', genit. *devyās*, and on the other by Vedic *vr̥kīs* 'she-wolf', genit. *vr̥kyās* (which spells *vr̥kīas*). The *devī*-type commonly forms feminines to athematic adjectives such as *nt*-participles; examples include Ved. *yuñjat-ī* 'yoking (fem.)' (< **jung-nt-ih₂*) and Gk. *phérousa* 'carrying (fem.)' (< pre-Greek **bher-ont-ja*). (On the different syllabifications of **-ih₂*, see §3.42.) There is no sure evidence of either of these types in Anatolian.

There are scattered vestiges of other feminine formations, of which we shall mention one. A feminine suffix **-s(o)r-*, derived from a word **soro-* 'woman', is indicated by the Indo-Iranian and Celtic feminine of the numerals 'three' and 'four', **t(r)i-sr-* and **k^wetur-sr-* (Ved. *tisrās*, *cātasras*; Av. *tišrō*, *catagrō*, with *-gr-* from **-sr-*; Gaul. *tides*; and OIr. *teoir*, *cetheoir*). It is also found as a feminizing suffix for certain Hittite nouns (see §9.44a), and may well be the second element in such kinship terms as **sue-sōr* 'sister', perhaps 'woman of one's own (kinship group)' (**suo-* 'one's own'; see §7.12).

Adjectives

6.72. Indo-European adjectives did not differ from nouns in declension, and could be thematic or athematic. To judge by the older IE languages, any adjective could also be used as a noun (like English *the good*, *the bad*, and *the ugly*). Sometimes the substantivization of an adjective (that is, the process of turning it into a noun) involved morphophonological changes, especially in an idiomatic meaning. Among the best-known of these is retraction of the accent when turning an adjective into a noun, as in Ved. *kṛṣṇá-* 'black' versus the noun *kṛṣṇa-* 'black antelope', and Gk. *dolíkhós* 'long' versus *dolíkhos* 'long race-course'.

Formation of adjectives

6.73. Adjectives could be formed by the addition of suffixes to roots or word-stems. While most of the adjectives so formed were *o*-stems (see the list starting in §6.74 below), many other adjectives were athematic. For example, a number of adjectives that were formed from roots of adjectival meaning were *u*-stems (as **sueh₂du-* 'sweet' > Vedic *svādú-*, Doric Gk. *hādús*), and we have seen various consonant-stem participles like those in **-nt-* and the feminine adjectives in **-ih₂-*. A type of *i*-stem adjective of some interest is that exemplified by Latin *prō-clīu-is* 'steep' and *dē-pil-is* 'hairless': each of these is a compound formed from an *o*-stem noun (*clīu-us* 'a slope', *pilus* 'hair'), with the thematic declension replaced by *i*-stem endings when the noun formed a compound adjective. Aside from suffixation, athematic adjectives could be formed from athematic nouns by internal derivation, detailed in §6.29.

Adjectival suffixes

6.74. **-jo-* and **-jō-*. Indo-Europeanists have traditionally reconstructed an all-purpose adjectival suffix **-jo-* underlying such forms as Hitt. *ištarniya-* 'central'

(< *īstarna* 'between'), Ved. *dāmiya-* 'domestic' (< *dām-* 'home'), *gāvya-* 'pertaining to cows' (< *gāu-* 'cow'), and Lat. *ēgregius* 'outstanding' (< *ē grege* 'out of the herd'). But recent research suggests that PIE originally had two or more similar but functionally distinct suffixes that merged to a monosyllabic **-iō-* and a disyllabic **-iō-*. According to one widely held view, locational adjectives like *dāmiya-* were originally possessive derivatives formed by adding **-o-* (or **-ō-*, §6.60) to locatives in **-i-*; the resultant sequence **-i-o-* tended to be phonetically realized as **-iō-*. Other sources have been forwarded as well; in all of them, **-o-* was added to some type of stem containing an *-i-*. The issue is made more difficult because **-iō-* could come to be pronounced **-iō-* in individual branches, resulting in variant forms of the same adjective: Gk. *mésos* and Gothic *midjis* point to **medhjos* as the PIE adjective for 'middle', whereas Latin *medius* presupposes **medhjos* (as **medhjos* would have become **meius* by Latin sound laws). For ease of presentation, we will simply use the traditional reconstruction **-iō-* in most cases.

Aside from being freely added to other parts of speech for forming adjectives, **-iō-* was very commonly employed as a compositional suffix, as in Lat. *ēgregius* above. In several branches, the suffix was used to form patronymics ('son of . . .'), e.g. Homeric Gk. *Telamōn-íos Aías* 'Aias (Ajax) son of Telamon', Old Persian *Haxāmaniš-īya-* 'offspring of Achaemenes, Achaemenid', Lat. *Seru-ius* 'son of a slave (*seruus*), Servius'.

6.75. **-ko-*. This suffix is most commonly found added to nouns to indicate origin or material composition: Ved. *sindhu-ka-* 'from Sindh', Gk. *Libu-kós* 'Libyan', Gaulish (Latinized) *Are-mori-cī* 'those by the sea, Aremoricans', Goth. *staina-hs* 'stony'. It frequently appeared in a longer version **-iko-* as a suffix indicating appurtenance, as in Gk. *hipp-ikós* 'having to do with horses', Lat. *bell-icus* 'pertaining to war', and OE *gyd-ig* 'dizzy' (< Germanic **gud-iga-* 'possessed by a god'; Mod. Eng. *giddy*).

Related is **-isko-*, found in Germanic and Balto-Slavic to indicate affiliation or place of origin: OHG *diut-isc* 'pertaining to the (common) people' (> Mod. German *Deutsch* 'German'), OCS *rām-iskŭ* 'Roman', and Russ. *russkij* 'Russian'. This is the source of Mod. Eng. *-ish*.

6.76. **-ro-*. This suffix took part in the derivational system known as the Caland system (see §6.87). It was added to the zero-grade of an adjectival root to form that root's free-standing adjectival form, and was usually accented, as in **h₁rudh-ró-* 'red' (Ved. *rudhirá-* 'blood-red, bloody', Gk. *eruth-rós*, Lat. *ruber*) and **h₂ṛǵ-ró-* 'swift; shining' (Ved. *ṛj-rá-*, Gk. *argós* < **argrós* 'swift').

6.77. **-tó-*. Aside from forming passive verbal adjectives (see §5.57), this suffix formed possessive adjectives, as in Lat. *barbā-tus* 'bearded', OCS *bogatŭ* 'wealthy' (from *bogŭ* 'wealth'), and Eng. *beard-ed*. For (unaccented?) **-to-* in superlatives and ordinals, see §6.81 and §§7.19, 7.21.

Comparison of adjectives

COMPARATIVE

6.78. There were two comparative suffixes in PIE. The first was an amphikinetic ablauting suffix **-jos-* (masculine nomin. **-jōs*, weak cases **-is-*) added to the full grade of the root: Av. accus. sing. *nāud-iiāgham* 'weaker' (< Indo-Iranian **nādh-jās-am*); Mycenaean Gk. *me-zo-e* '(two) bigger ones' (< pre-Greek **meg-jos-e*); Lat.

maior (Archaic Lat. *maiōr*), neut. *maius* (**mag-īos-*), adverb *magis* 'more'; OIr. *sin-iu* 'older' (**sen-īōs*, cp. Lat. *senior*); Goth. *manag-iz-* 'more'; and OCS *bol-īš-* 'bigger'.

6.79. The second comparative suffix was *-*tero-*. This was not originally a comparative suffix: its basic function was to set off a member of a pair contrastively from the other member of the pair, as in Gk. *skaiós* 'left' vs. *deksiterós* 'right', i.e. 'the one on the right (as opposed to the one on the left)'. Thus a comparative like Gk. *sophō-teros* 'wiser' meant originally 'the wise one (of two people)', hence '(the one that is) more wise'. This suffix is continued by the comparative suffixes -*tara-* of Sanskrit and -*tero-* of Greek, and by such forms as Lat. *dē-ter-ior* 'farther down, lower, worse', OIr. *úachtar* 'higher part', and Eng. *far-ther*.

SUPERLATIVE

6.80. To judge by the fact that some of the older IE languages could use the positive degree of an adjective plus a genitive plural noun to express the superlative of the adjective, PIE did not necessarily express the superlative degree by any change to the adjective. Thus Homeric Greek has *dīa theāōn* (genit. pl.) 'divine of/among goddesses = most divine goddess', and Hittite has *šallayaš* (genit. pl.) *šīunaš* (genit. pl.) *šalliš* 'great among the great gods = the greatest of the great gods'.

6.81. However, at least two suffixes are reconstructible that have a superlative function in the daughter languages. One is *-*mmo-*, seen in various forms and combinations: alone in Lat. *min-imus* 'least, smallest'; *-*t-mmo-* in the Skt. superlative suffix -*tama-*; and *-*is-mmo-* (a combination of the weak stem *-*is-* of the comparative suffix and the superlative *-*mmo-*) in Latin superlatives like *fort-issimus* 'strongest' and *maximus* 'biggest' (probably **magis-mmo-*). The other is the addition of *-*to-* to the zero-grade *-*is-* of the comparative suffix, yielding *-*isto-*, as found in the Skt. superlative suffix -*iṣṭha-*, Gk. *még-istos* 'greatest', and Goth. *bat-ista* 'best' (and continued by the Eng. superlative suffix -*est*). These superlative suffixes bear a striking resemblance to, and are perhaps ultimately identical with, certain ordinal numeral suffixes, e.g. *-*isto-* in Eng. *first*; see §§7.18ff.

Nominal Composition and Other Topics

6.82. PIE possessed a rich set of compound noun formations. Compounds can be broadly classified into two types according to their meaning. In one type, which can be illustrated by *blackbird* ('a type of) black bird', the compound is essentially the sum of its parts, and its referent (a type of bird) is one of the compound members itself (usually the second one, as here). This type is called *endocentric* (or determinative). Examples of endocentric compounds from other IE languages include Skt. *Siṃha-puras* 'Lion City' (*Singapore*), Gaulish (Latinized) *Nouio-dūnum* 'New Fort', German *Blut-wurst* 'blood sausage', and OCS *dobro-dějanje* 'good works'.

The second type, which can be illustrated by another bird-name, *redthroat*, is more than the sum of its parts and refers to something outside itself: the referent is not a type of throat, but a type of bird possessing a red throat. Such compounds are called *exocentric* or possessive compounds, but in Indo-European circles they are typically called *bahuvrihis*, a term borrowed from Sanskrit (*bahuvrihi-*, literally

'much-riced', i.e. 'having much rice' – itself an example of the type). Bahuvrihis are quite common in personal names: Gk. *Aristó-dēmos* 'best-people = having the best people', OIr. *Fer-gus* 'hero-strength = having a hero's strength', Cz. *Bohu-slav* 'having the fame of god', Lepontic *Uvamo-kozis* 'having supreme guests', and Eng. *Pea-body*. (Recall §2.48.)

6.83. It is frequently said that bahuvrihis typically have o-grade of the ablauting syllable of the second compound member. Such is indeed the case in such forms as Gk. *eu-pátor* 'having a good father' (from *patēr* 'father') and Umbrian *du-purs-* 'having two feet' (< **-pod-*). These o-grades are likely to be survivals of old amphikinetic inflection rather than engendered directly by the process of compounding; but the matter is in need of further investigation.

6.84. Some bahuvrihis had adverbs as the first compound member: Gk. *én-theos* 'having god within = inspired', Lat. *dē-plūmis* '(with) the feathers away = featherless', Gaul. *exs-ops* '(with) the eyes away = blind'.

6.85. PIE also had compounds in which one member was a verb or a verbal derivative that governed the other member. Having the verb as second compound member (type Eng. *bar-keep*, *cow-herd*) is quite common, e.g. Ved. *rathe-ṣṭhā-* 'standing on a chariot, chariot-fighter', Lat. *armi-ger* 'one who bears arms', Arm. *ban-ber* 'word-carrier, messenger', and OCS *medv-ědi* 'honey-eater = bear'. More rarely, the verb is the first member (type Eng. *pick-pocket*), as in Gk. *phérē-oikos* 'carry-house = snail' and OCS *vě-glasŭ* 'know-voice = knowledgeable'.

6.86. Some adjectives and other parts of speech appear in a reduced form as prefixes. Some well-known examples include the zero-grade prefix **h₁su-* from the adjective **h₁esu-* 'good' (Skt. *su-*, Gk. *eu-*, OIr. *so-*, OCS *sŭ-*); the so-called "privative" or negating prefix **h₂-* 'not, un-' from **ne* 'not' (see §7.24); and the numeral prefix **tri-* 'three-' from **tréies* 'three' (see §7.17).

The Caland system

6.87. The nineteenth-century Sanskritist Willem Caland noticed that adjectives in *-*ro-* (§6.76) substituted an *-*i-* for that suffix when they were the first member of a compound. For example, Av. *dərəzra-* 'strong' becomes *dərəzi-* in compounds like *dərəzi-raθa-* 'having a strong chariot', and Gk. *kūdrós* 'glorious' exists alongside *kūdi-áneira* 'bringing glory to men'. This rule of morphological replacement became known as "Caland's Law," and since Caland's time it has been discovered that several other suffixes participate in a so-called Caland system. Chief among these suffixes, besides *-*ro-*, are the stative suffix *-*eh₁-* and its relatives (§5.34), the adjectival suffix *-*u-*, and several others. The roots involved are typically adjectival.

For Further Reading

As with the verb, no comprehensive treatment of the noun has been written in nearly a century. The accent-ablaut classes are the subject of many articles; the seminal ones are mostly in German by such authors as Jochem Schindler and Heiner Eichner, building on work by earlier scholars. Rieken 1999 is a massive study of Hittite athematic nouns and has an excellent detailed overview of the athematic accent-ablaut types. Nussbaum 1986 is a superb lexical and morphological study, treating many issues in the derivational relationships

among IE nominal formations. Many of Nussbaum's articles concern nominal morphology and derivation and can be read with profit.

For Review

Know the meaning or significance of the following:

nominative	directive	endingless locative	heteroclitic
vocative	collective	root noun	r/n-stem
accusative	strong case	accent-ablaut class	Hoffmann suffix
genitive	weak case	acrostatic	<i>tomós/tómos</i> nouns
ablative	animate	proterokinetic	vrdhī-derivative
dative	masculine	amphikinetic	bahuvrīhi
instrumental	feminine	hysterokinetic	Caland system
locative	neuter	internal derivation	

Exercises

- 1 The following are cases of the athematic noun **pod-* (**ped-*) 'foot'. Identify the case and number of each form:

a **pód-ŋ* b **ped-bhī(-)* c **ped-ōm* d **ped-ŋs*

- 2 Below are given various case-forms of the animate *o*-stem noun **logos* 'word'. Identify the case and the number of each form. Some forms have more than one answer.

a **loḡosjo* c **loḡōs* e **loḡē* g **loḡoh₁*
b **loḡōns* d **loḡos* f **loḡom* h **loḡōt*

- 3 For the neuter *o*-stem noun **deh₃nom* 'gift', give each form indicated below:

a nominative plural b accusative singular c dative singular

- 4 Using the information in §§6.45ff., give the full PIE singular and plural paradigms of the nouns in 2 and 3 above.

- 5 Below are given the nominative, accusative, and genitive singular of a variety of PIE athematic nouns. Identify the accent-ablaut class to which each noun belongs. Specify if the noun is a root noun.

a **péh₃-ti-s* **péh₃-ti-m* **ph₃-tei-s* 'a drink'
b **dhéh₁-mṇ* **dhéh₁-mṇ* **dh₁-mén-s* 'something placed'
c **-stéh₂-s* **-stéh₂-ŋ* **-sth₂-és* 'standing on ...'
d **h₂st-ér* **h₂st-ér-ŋ* **h₂st-r-és* 'star'
e **djéu-s* **djéu-m* **diu-és* 'sky, sky-god'
f **dh₃-tér* **dh₃-tér-ŋ* **dh₃-tr-és* 'giver'

6 Below are a set of nouns derived from verbs. Provide a translation of the nouns.

- | | | | |
|---|---|---|--|
| a | * <i>ġneh</i> ₃ - 'know': * <i>ġneh</i> ₃ - <i>mġ</i> | d | * <i>g^when</i> - 'kill': * <i>g^whon</i> - <i>ós</i> |
| b | * <i>ġenh</i> ₁ - 'beget, create': * <i>ġenh</i> ₁ - <i>tör</i> | e | * <i>peh</i> ₃ - 'drink': * <i>peh</i> ₃ - <i>tlom</i> |
| c | * <i>uék^w</i> - 'speak': * <i>uék^w</i> - <i>os</i> , * <i>uék^w</i> - <i>es</i> | | |

7 Form the indicated derivative of the noun or adjective below.

- a **k^wtyr-ped*- 'having four feet': feminine
 b **h₂rtkos* 'bear': feminine
 c **smókġr* 'beard': athematic adj. 'having a beard'

8 Mention was made above (§6.73) of *u*-stem adjectives formed from roots having adjectival meaning. The suffix *-*ro*- was also frequently used in this purpose. The distribution of these suffixes seems to have been phonologically conditioned by the sounds in the root. Based on the forms below, what restriction was placed on the kinds of roots to which *-*u*- could be attached?

- **h₂reudh*- 'red': adjective **h₁rudh-ro*-
 **sueh₂d*- 'sweet': adjective **sueh₂d-u*-
 **ók*- 'swift': adjective **ók-u*-
 **keud*- 'glorious': adjective **kud-ro*-
 **mreġh*- 'short': adjective **mġġh-u*-
 **h₁leng^wh*- 'light, quick': adjective **h₁lng^wh-u*-
 **gheu*- 'be in awe of': adjective **gheu-ro*- 'causing awe'
 **bherġh*- 'high': adjective **bhrġh-u*-
 **deuh₂*- 'long in duration': adjective **duh₂-ro*-
 **angh*- 'tight, constricted': adjective **angh-u*-

9 Identify each of the following English compounds as either endocentric or exocentric.

- | | | | | | |
|---|------------------|---|------------------|---|--------------------|
| a | <i>redhead</i> | c | <i>blueberry</i> | e | <i>blackboard</i> |
| b | <i>greenback</i> | d | <i>whitecap</i> | f | <i>yellowbelly</i> |

7 Pronouns and Other Parts of Speech

Pronouns: Introduction

7.1. Pronouns present special difficulties to the comparatist. First of all, they easily come under analogical influence of each other. This is true especially of the personal pronouns: the accusative singulars of the first and second personal pronouns came to rhyme not only in English (*me thee*) but also in Hittite (*ammuk tuk*), Sanskrit (*mām tvām*), Greek (*emé sé*), and Russian (*menja tebja*), among many other languages; yet none of these formations is precisely cognate, and each is found only in its own branch of the family. Second, pronouns tend to attract ancillary emphatic particles or affixes of various kinds to give them greater saliency. This is true especially of deictic and demonstrative pronouns, as in dialectal English *this-(h)ere*, *that-(th)ere* replacing simple *this* and *that*. Third, pronouns tend to have a broad range of idiomatic uses that often make reconstruction of their original functions uncertain.

In spite of these difficulties, we are fairly well informed about the PIE pronominal system. We can reconstruct personal, reflexive, demonstrative, relative, and interrogative pronouns for the proto-language. We know that, like adjectives, pronouns were declined for case and number and (except for the personal and reflexive pronouns) for gender; and we know that the pronominal declensions differed from the nominal and adjectival declensions in several respects discussed below.

Personal Pronouns

7.2. The personal pronouns of the IE languages are of a somewhat paradoxical historical nature: probably no other set of words contains elements that are as ancient (they have been termed the “Devonian rocks” of the proto-language), yet at the same time they have undergone a great deal of analogical remodeling. Everywhere they are exceptional in the context of their language’s grammar. In English, only the personal pronouns distinguish a subjective from an objective case (as in *I* vs. *me* and *she* vs. *her*), a retention of the fuller case-system that English once had. Going back a millennium, the Old English personal pronouns were the only words to have a set of dual forms alongside singular and plural – another archaism at the time. The use of more than one stem in any given paradigm (such as *I/me, we/us*), a phenomenon called *suppletion*, recurs throughout the family – and with exactly the same stems; the suppletion is reconstructible for PIE and harks back to extremely old layers of the language’s morphology.

But the amount of analogical refashioning that has occurred makes it impossible to reconstruct most pronominal forms with any certainty, aside from the bare stems. We know that PIE had complete singular, dual, and plural paradigms for the first and second personal pronouns, and had different paradigms for stressed and unstressed (enclitic) forms. We also know that PIE had no special forms for the third person, probably using demonstrative pronouns instead, such as *so **seh₂* **tod* (see §7.9). There was also a reflexive pronoun that was indifferent to both person and number.

First person

7.3. A sampling of the comparative data for the first personal pronoun, in singular, dual, and plural, is as follows:

	Hitt.	Gk.	Ved.	Av. ³	Lat.	Goth.	Arm.	Toch. B	OCS	Lith.
sg.										
nom.	<i>n̄k</i>	<i>egó(n)</i>	<i>ahám</i>	<i>axəm</i>	<i>ego</i>	<i>ik</i>	<i>es</i>	<i>ñāś</i>	<i>axp̄</i>	<i>àš</i>
acc.	<i>anmuk</i>	<i>emé</i>	<i>mám</i>	<i>mam</i>	<i>mē</i>	<i>muk</i>	<i>is</i>	<i>ñāś</i>	<i>mene</i>	<i>manė</i>
gen.	<i>ammēl</i>	<i>eméo¹</i>	<i>māma</i>	<i>mana</i>	<i>mei</i>	<i>mema</i>	<i>im</i>	<i>ñu</i>	<i>mene</i>	<i>manės</i>
dat.	<i>ammuk</i>	<i>emoi</i>	<i>māhyam</i>	<i>maibnā⁴</i>	<i>mabī</i>	<i>mīs</i>	<i>inj</i>		<i>mīnė</i>	<i>mān</i>
du.										
nom.		<i>nó</i>	<i>vām</i>			<i>unt</i>		<i>wene</i>	<i>vě</i>	<i>vėdu⁵</i>
acc.		<i>nó</i>	<i>dvām</i>	<i>ōdūma⁴</i>		<i>ugkis</i>		<i>wene</i>	<i>na</i>	<i>nuodu⁵</i>
gen.		<i>nōm</i>	<i>dvāyos</i>						<i>naju</i>	
dat.		<i>nōm</i>				<i>ugkis</i>			<i>nama</i>	
pl.										
nom.	<i>wēš</i>	<i>ámmeš²</i>	<i>vayām</i>	<i>vaēm</i>	<i>nōs</i>	<i>weis</i>	<i>mek^e</i>	<i>wes</i>	<i>my</i>	<i>mēs</i>
acc.	<i>anzāš</i>	<i>ámmeš²</i>	<i>asmān</i>	<i>āhmā⁴</i>	<i>nōs</i>	<i>uns(is)</i>	<i>mez</i>	<i>wes</i>	<i>ny</i>	<i>mūs</i>
gen.	<i>anzāl</i>	<i>ámmonš²</i>	<i>asmākam</i>	<i>ahmākam</i>	<i>nōbis</i>	<i>unsara</i>	<i>mer</i>	<i>wesi</i>	<i>nasū</i>	<i>mūsū</i>
dat.	<i>anzāš</i>	<i>ámmsi²</i>	<i>asmābhyam</i>	<i>ahmaibhiā⁴</i>	<i>nostrum</i>	<i>uns(is)</i>	<i>mez</i>		<i>namū</i>	<i>mūms</i>

¹ Homeric. ² Aeolic. ³ Younger Avestan except where noted. ⁴ Old Avestan. ⁵ Dialectal.

NOTES ON THE RECONSTRUCTION OF THE FIRST PERSON

7.4. For the nominative singular of the first person singular, we reconstruct **eǵoh₂* (extended to **eǵh₂-om* in e.g. Vedic *ahám*). The other cases are built from a different stem, **me-* (sometimes extended to **eme-*). The dual and plural contain at least two stems as well, one in **ue(i)-* or similar for the nominative, and another in **ns-* for the other cases. The genitive singular of the first person may have been **mene*, to judge by the Indo-Iranian and Balto-Slavic forms above, as well as by Welsh *fy* 'my', which once ended in a nasal since it nasalizes the following sound (as *fy nghar* 'my car'; nasalization will be discussed in chapter 14). A dative singular **me-ǵh(i)* is reflected in Indo-Iranian, Latin, and probably Armenian *inj* (< **eme-ǵh-*). The nominative dual is fairly securely reconstructible as **ueh₁* (> **uē*), although the vowel is short in Germanic and Baltic. The nominative plural was probably **uei*. The plural oblique cases are apparently ultimately based on a stem **ns-me-*; in some branches (Armenian, Baltic), only the second part of this seems to occur, perhaps under the influence of the 1st plural verbal endings in *(*)-me-*. The element **ns-* is the zero-grade of **nes*; this and other grades are found in the Latin pl. *nōs* and in a number of enclitic dual and plural forms.

Second person

7.5. A selection of the comparative data for the second person follows:

	Hitt.	Gk.	Ved.	Av.	Lat.	Goth.	Arm.	Toch. B	OCS	Lith.
sg.										
nom.	<i>zuk</i>	<i>tú¹</i>	<i>tvám</i>	<i>tā</i>	<i>tū</i>	<i>þu</i>	<i>dow</i>	<i>t(u)we</i>	<i>ty</i>	<i>tù</i>
acc.	<i>tuk</i>	<i>tē¹</i>	<i>tvám</i>	<i>θβam</i>	<i>tē</i>	<i>þuk</i>	<i>k'ez</i>	<i>ca</i>	<i>tebe</i>	<i>tavè</i>
gen.	<i>twēl</i>	<i>téos¹</i>	<i>táva</i>	<i>tauua</i>	<i>tuī</i>	<i>þeina</i>	<i>k'o</i>	<i>tañ</i>	<i>tebe</i>	<i>tavþs</i>
dat.	<i>tuk</i>	<i>toi¹</i>	<i>tūbhryam</i>	<i>taibiia³</i>	<i>tibi</i>	<i>þus</i>	<i>k'ez</i>		<i>tebē</i>	<i>táu</i>
du.										
nom.		<i>spbó</i>	<i>yuvám</i>					<i>yene</i>	<i>va</i>	
acc.		<i>spbó</i>	<i>yuvám</i>			<i>igqis</i>		<i>yene</i>	<i>va</i>	
gen.		<i>spbóm</i>	<i>yuvós</i>	<i>yauuākam</i>		<i>igqara</i>			<i>vaju</i>	
dat.		<i>spbóin</i>				<i>igqis</i>			<i>vama</i>	
pl.										
nom.	<i>šumēš</i>	<i>šummes²</i>	<i>yūyám</i>	<i>yūš¹</i>	<i>uūs</i>	<i>jūs</i>	<i>dowk¹</i>	<i>yes</i>	<i>vy</i>	<i>jūs</i>
acc.	<i>šumās</i>	<i>šumme²</i>	<i>yušmán</i>		<i>uōs</i>	<i>izwis</i>	<i>jez</i>	<i>yes</i>	<i>vasū</i>	<i>jūs</i>
gen.	<i>šumenzan</i>	<i>ummoñ²</i>	<i>yušmākam</i>	<i>yūšmākam</i>	<i>uestram</i>	<i>izwara</i>	<i>jer</i>	<i>yesı</i>	<i>vasū</i>	<i>jūsų</i>
dat.	<i>šumāš</i>	<i>ummu²</i>	<i>yušmābhryam</i>	<i>yūšmaouō</i>	<i>uōbis</i>	<i>izwis</i>	<i>jez</i>		<i>vamū</i>	<i>jūms</i>

¹ Doric. ² Aeolic. ³ Old Avestan.

RECONSTRUCTION OF THE SECOND PERSON

7.6. The singular forms of the second person are based on a stem in **tu-*. The dual and plural forms show mostly stems in **us-* and **ju-*. The nominative singular was perhaps **tuH*, with the laryngeal the source of the long vowel in most of the descendant forms. The oblique plural stem is usually reconstructed as **us-me-*, but it has been recently suggested by the American Indo-Europeanist Joshua Katz that this was remade (under the influence of the first person **us-me-*) from **us-me-*, reflected in Goth. *izwis* and related forms.

Enclitic personal pronouns

7.7. A number of the branches, in particular Anatolian, Indo-Iranian, Greek, Balto-Slavic, and Tocharian, distinguish fully stressed (emphatic or contrastive) personal pronouns from unstressed clitic object pronouns. For the dative of the first singular pronoun, for example, Hittite has the stressed form *ammuk* and clitic *-mu*, Greek has *emoi* and clitic *moi*, Vedic has *māhyam* and clitic *me*, and Lithuanian has *mán* and (Old Lith.) clitic *-m(i)*. Observe the forms below:

	1st person					2nd person				
	Hitt.	Gk.	Ved.	Av.	OCS	Hitt.	Gk.	Ved.	Av.	OCS
sg.										
acc.	<i>-mu</i>	<i>me</i>	<i>mā</i>	<i>mā</i>	<i>mę</i>	<i>-ta, -du⁴</i>	<i>te</i>	<i>tvā</i>	<i>θβā</i>	<i>tę</i>
dat.	<i>-mu</i>	<i>moi</i>	<i>me</i>	<i>mōi,² mē</i>	<i>mi</i>	<i>-ta, -du⁴</i>	<i>toi</i>	<i>te</i>	<i>tōi,² tē</i>	<i>ti</i>
gen.	<i>-mu</i>	<i>meu¹</i>	<i>me</i>	<i>mōi,² mē</i>		<i>-ta, -du⁴</i>	<i>seu¹</i>	<i>te</i>	<i>tōi,² tē</i>	
du.										
A.D.G.			<i>nau</i>					<i>vām</i>		<i>va</i>

1st person					2nd person				
Hitt.	Gk.	Ved.	Av.	OCS	Hitt.	Gk.	Ved.	Av.	OCS
pl.									
A.D.G.	- <i>naš</i>	<i>nas</i>	<i>nā³</i> , <i>nō</i>	<i>ny</i>	- <i>šmaš</i>		<i>vas</i>	<i>vō</i>	<i>vy</i>
abl.			<i>vā²</i>						

¹ Homeric. ² Old Avestan. ³ Old Avestan accusative. ⁴ Used when together with the reflexive particle *-za*.

Not surprisingly, the clitic pronouns are mostly just reduced versions of the fuller forms. One can reconstruct for the singular of both persons **me* and **te* in the accusative and **moi* and **toi* in the dative and genitive (Gk. *meu* and *seu* are innovations); and for the plural in all oblique cases, **nos* and **uos*.

Anatolian, alone among the branches, has innovated a set of third-person clitic subject pronouns (e.g. Hitt. *-aš* 'he', *-e* 'they' [animate]). They are used only with certain kinds of intransitive verbs, as will be discussed in §9.11.

Other Pronouns and the Pronominal Declension

7.8. Among other PIE pronouns, the most easily reconstructible are the demonstratives in **so-*, **to-*, and **ei-*, the relative **(H)jo-*, the indefinite/interrogative **kʷo-*, and the reflexive **sme-* (all discussed below).

Several of the case-endings differ from those of the nominal and adjectival declensions. The neuter nominative-accusative singular ended not in **-m* but in **-d* (**tod*, **kʷid*, etc.; also reconstructed as **tot*, **kʷit*; cf. §3.40). The genitive singular ending was **-eso*, as in OCS *česo* 'of what' and Goth. *þis* 'of the'. (In Germanic, this ending spread to the thematic noun declension: Goth. *dagis* 'of a day'.) The masculine nominative plural ending was **-oi* (so **toi*, **kʷoi*, etc.), an ending that spread to the thematic noun declensions in several branches (§6.43). Finally, a genitive plural ending **-sōm* has been reconstructed on the basis of Hitt. *appen-zan* 'of those', Ved. *e-śām* 'of these', and Celtiberian *soi-sum* 'of these'.

A formant **-sm-* is found in several oblique cases in some of the languages: Ved. *asmin* and South Picene *esmen*, both locative singulars meaning 'in this'; OPruss. dative sing. *schismu* '(for) this', and probably Goth. dative sing. *þamma* 'for the, for this'. Surely related are forms without the *-s-* such as Lith. locative *tamè* 'in that'.

An additional peculiarity of the pronominal declensions is that, though basically thematic, they often show *i*-stem forms. Thus the interrogative pronoun **kʷo-* had nominative/accusative forms in **kʷi-* such as Av. *ciš* 'who?', Gk. *tis* 'who?' (in this language the *i*-stem forms were generalized throughout the paradigm), and Lat. *quid* 'what?' A stem **ki-* 'this' (see next section) is an *i*-stem, although the related emphatic particle **ke* (§7.28) ends in an *-e* that presupposes thematic forms as well.

Demonstrative pronouns

7.9. Widely represented are the two stems **so-* and **to-* meaning 'this, that' or 'the'. The first of these seems to have been marginalized early, and was incorporated already in PIE into the paradigm of **to-*, for which it formed the nominative masculine and feminine. We can thus reconstruct a pronoun with the nominatives **so* (masc.),

**seh*₂ (fem.), **tod* (neut.), becoming Ved. *sá sá tát*, Av. *hō hā tat*, Gk. *hō hē tó*, Goth. *sa so þata*, and Toch. B *se sā te*. Another stem **ei*- 'this' is reflected for example in Skt. *ay-ām* (masc.) *id-ām* (neut.); Av. *im* (accus.) 'him'; Lat. *is ea id* 'this; he, she, it'; and Goth. *is* 'he'. A stem **kī*- 'this' is reflected principally in Balto-Slavic and Germanic (as in OCS *sī* 'this', Lith. *šis* 'this', and Eng. *he*, and OHG *hiu-tagu* 'on this day, today' [> modern German *heute*]), and vestigially in such forms as Lat. *cis* 'on this side of'. Another pronominal stem, **eno*- or **ono*-, probably extensions of a stem **e*-, is found in Hitt. *ēni*- 'that one' and OCS *onŭ* 'that'.

Relative pronoun

7.10. Securely reconstructible is a relative pronominal stem **io*- (or **Hio*-; the presence of the laryngeal is disputed, and depends on one's interpretation of the initial *h*- of Greek *hó*-, cp. §12.22): Ved. *yā*-, Av. *ya*-, Gk. *hó*-, and Celtic **yo*- (in Gaulish *dugiōnti-io* 'who serve'). Other branches have marshalled the indefinite/interrogative stem **kʷo*- (next section) into service as the relative pronoun.

Interrogative/indefinite pronoun

7.11. A stem **kʷo*- is widely attested, functioning both as an interrogative ('who, what?') and as an indefinite ('someone, something'): Ved. *kás* and Av. *kō* 'who?', Gk. adverbial interrogatives and indefinites in *po*- (e.g. *poŭ* 'where?', *pōs* 'somehow'), Goth. *hvas* 'who?', OCS *kŭ-to* 'who?', and Lith. *kàs* 'who?'. The *i*-stem variant **kʷi*- is reflected in Hitt. and Luv. *kuiš* 'who?', Gk. *tīs* 'who?', Lat. *quis* 'who?', OIr. *cia* 'who?', OE *hwī* 'how, why?', and OCS *čŭ-to* 'what?'. In the meaning 'anyone' or as an indefinite relative 'whoever' it is often doubled, as in Hitt. *kuiš kuiš*, Lat. *quisquis*, and Oscan *pitpit* 'whatever'. In several branches (Anatolian, Italic, Germanic, and Balto-Slavic), this stem is used for the relative pronoun, as in Eng. *which*.

Reflexive pronoun

7.12. A reflexive pronoun **sue*- (also **se*-) meaning '(one)self' has descendants in numerous branches: Ved. *sva-yám*, Av. *xʷāi* (dat.), Gk. (Pamphylian) *whe* (Classical Gk. *hé*), Lat. (accus.) *sē*, OIr. *fa(-dessin)*, and OCS (accus.) *se*. It did not have a nominative case, did not distinguish number, and could be used with any of the three persons. The stem also formed the basis for the reflexive adjective **syo*- meaning '(one's) own', reflected in Ved. *svá*-, Gk. *heós*, Lat. *suus*, and OCS *svojŭ*. The syntax of the reflexive pronoun and adjective will be discussed in §8.18.

Pronominal adjectives

7.13. The best-attested pronominal adjective is **al-jo*- 'other' in Gk. *állos*, Lat. *alius*, Goth. *aljis*, OIr. *aíl*, and Toch. B *alyek*.

Numerals

7.14. Numerals are technically adjectives or quantifiers, but morphologically they stand somewhat apart from other adjectives. The cardinal numerals were indeclinable except for 1–4, which not only declined but also distinguished gender.

7.15. No single form for the number 'one' can be reconstructed; there were at least two roots for the concept. The most widely represented is *oi-, variously suffixed, especially as *oi-no- (Lat. *ūnus*, OIr. *óen*, Goth. *ains* [Eng. *one*], OCS *inŭ*, and cp. Gk. *oinós* 'roll of one in dice'); *oi-ko- and *oi-yo- are also found (e.g. Ved. *éka-*, Av. *aēnuua-*). The other root was *sem-, the base of Gk. *heís* (*sem-s), feminine *mía* (*sm-ih₂), Arm. *mí*, and Toch. A *sas*, B *še*. This root fundamentally expressed identity; it is the root of Eng. *same* and Lat. *similis* 'similar, like'.

7.16. A representative sampling of the cardinal numbers 2–10, 20, and 100 is given below. Blank spaces indicate that the relevant forms are not useful for reconstruction.

	Gk.	Ved.	Av.	Lat.	Welsh	Goth.	Arm.	Toch. A	OCS	Lith.
2	<i>duō</i>	<i>dvā(u)</i>	<i>dwua</i>	<i>duō</i>	<i>dau</i>	<i>twai</i>	<i>erkow</i>	<i>wu</i>	<i>diŭva</i>	<i>dù</i>
3	<i>treís</i>	<i>tráyas</i>	<i>θrāiō</i>	<i>trēs</i>	<i>tri</i>	<i>þreis</i>	<i>erek'</i>	<i>tre</i>	<i>trije</i>	<i>trys</i>
4	<i>téttares</i>	<i>catváras</i>	<i>caθwārō</i>	<i>quattuor</i>	<i>pedwar</i>	<i>fidwor</i>	<i>čork'</i>	<i>štwar</i>	<i>četyre</i>	<i>keturi</i>
5	<i>pénte</i>	<i>pāñca</i>	<i>panca</i>	<i>quinque</i>	<i>pump</i>	<i>fimf</i>	<i>hung</i>	<i>pāñ</i>	<i>peti</i>	<i>penki</i>
6	<i>héks</i>	<i>ṣáṭ</i>	<i>xšumaš</i>	<i>sex</i>	<i>chwech</i>	<i>saihs</i>	<i>vec'</i>	<i>pāk</i>	<i>šesti</i>	<i>šeši</i>
7	<i>heptá</i>	<i>saptá</i>	<i>hapta</i>	<i>septem</i>	<i>saitb</i>	<i>sibun</i>	<i>ewi'n</i>	<i>špāt</i>	<i>sedmī</i>	<i>septyni</i>
8	<i>oktō</i>	<i>aṣṭā(u)</i>	<i>ašta</i>	<i>octō</i>	<i>wyth</i>	<i>abtan</i>	<i>oww'</i>	<i>okāt</i>	<i>osmī</i>	<i>aštuoni</i>
9	<i>enné(w)a</i>	<i>náwa</i>	<i>nanwa</i>	<i>novem</i>	<i>naw</i>	<i>nium</i>	<i>mn</i>	<i>nū</i>	<i>devęti</i>	<i>devyni</i>
10	<i>déka</i>	<i>dāśa</i>	<i>dasa</i>	<i>decem</i>	<i>deg</i>	<i>taihun</i>	<i>tasn</i>	<i>šāk</i>	<i>desęti</i>	<i>dėšimt</i>
20	<i>wíkati</i> ¹	<i>vimśati-</i>	<i>visati</i>	<i>uiginti</i>	<i>ugem(t)</i> ²		<i>k'san</i>	<i>wiki</i>		
100	<i>hekatón</i>	<i>śatām</i>	<i>satəm</i>	<i>centum</i>	<i>cant</i>	<i>hund</i>		<i>kānt</i>	<i>sūto</i>	<i>šimtas</i>

¹ Doric. ² Middle Welsh.

Based on these and some other data, we can reconstruct for the numerals 2–10 *d(u)hōh₁, *tréjes, *k^wétuores, *pénk^we, *syéks, *septri₁, *oktō(u), *néu₁, and *dék₁; for 20 *yíkmti (see below), and for 100 *kmtóm. It is widely thought that the words for 20 and 100 are derivatives of the word for 10. Under this view, *yíkmti is a simplification of a dual *dyih₁-dkmt-ih₁ 'two tens', and *kmtóm is simplified from an earlier *dkmtóm. Strictly speaking, the comparative evidence only supports a reconstruction *yíkmti, but most handbooks give *yíkmti because of the above analysis.

The numeral systems of the Anatolian languages are unfortunately almost entirely unknown. We only have Hieroglyphic Luwian *tuwi-* 'two' and Hitt. *teriyas* 'three' (genit. pl.), and maybe a *šiptam- 'seven' in Hitt. *šiptamiya-*, the name of a drink (perhaps containing seven ingredients; cp. Eng. *punch*, from Hindi *pañj* 'five').

As discussed in §6.71, the numerals '3' and '4' had special forms for the feminine with a suffix *-sr-, as in Ved. *tisrás* '3' and *cātasras* '4'.

Numerals in composition

7.17. When combined with other words as prefixes, typically to form bahuvrihis (like *three-toed* in English), special forms of the numerals are sometimes found. The most securely reconstructible are: *dyi- 'two-' (as in Ved. *dvi-pád-* 'two-footed', Gk. *dí-pod-* 'two-footed', Archaic Lat. *dvi-dent-* [Classical *bi-dent-*] 'sacrificial animal' having two teeth'); *tri- 'three-' (as in Ved. *tri-pád-* 'three-footed', Gk. *tri-pod-* 'three-footed [table]', Lat. *tri-ped-* 'three-footed', Gaulish *tri-garanus* 'having three cranes'); and *k^w(e)tru- or *k^wet₁- 'four-' (as in Ved. *cātus-pad-* 'four-footed', Av.

caθru-gaoša- 'four-eared', Gk. *tetrá-pod-* 'four-footed', and Lat. *quadru-ped-* 'four-footed'). The zero-grade of **sem-* 'one', **sm-*, is found as a prefix meaning 'one-, as one, together, same': Ved. *sa-kṛt* 'once', Gk. *há-ploos* 'one-fold, simple', *a-delphós* 'brother' (probably < **sm-g^welbh-o-* '[the brother] having one [= the same] womb, uterine brother'), and Lat. *sim-plex* 'one-fold'.

Ordinal numerals

7.18. The ordinal numerals cannot be reconstructed with precision because of the great variety of formations exhibited by the daughter languages.

The ordinal 'first' was formed from a base **prh₃-*, the zero-grade of a root **perh₃-* that is found in various adverbs and prepositions meaning 'forth, forward, front'; it is related to **prō* 'forth' (see §7.25 below). Widely represented are extensions of a stem **prh₃-mo-* (as in Ved. *pūrvīyá-* and OCS *prǐvŭ*) and a stem **prh₃-mo-* (e.g. Goth. *fruma*, Lith. *pirmas*; compare also Lat. *primus* from **pris-mo-*). Eng. *first* is from **prh₃-isto-*.

No form for 'second' can be reconstructed. PIE may not have used a derivative of 'two', given that the daughters typically use unrelated expressions, such as ones historically meaning 'the other' (e.g. OIr. *ail*) or 'the following' (e.g. Lat. *secundus*).

7.19. The ordinals from 'third' through 'sixth' were probably formed with the suffixes **-t-* or **-to-*: **tr₃-t-* or **tri₃-t-* 'third' (as in Ved. *trītiya-*, Gk. *trítos*, and Lat. *tertius*); **k^wetuy₃-t-* or similar for 'fourth' (as in Gk. *tétartos*, OE *feorþa*, and OCS *četvrtŭ*); **penk^w-to-* 'fifth' (as in Gk. *pémpptos* and Av. *puxda-*); and **suek(s)-to-* 'sixth' (as in Gk. *héktos* and Lat. *sextus*).

7.20. The ordinals 'seventh' through 'tenth' were, it seems, originally created simply by suffixing the thematic vowel **-o-* to the cardinal. This is most clearly seen in 'eighth', which was **oktōy₃-o-* or similar (as in Gk. *ógdo(w)os* and Lat. *octāuus*). In the case of 'seventh' and 'tenth', the cardinals ended in syllabic **-m*, and the combination **-m-o-* was realized as **-mmo-* phonetically: **septmmo-* (as in Gk. *hébdomos*, Lat. *septimus*, and OCS *sedmŭ*) and **dekmmo-* (as in Ved. *daśamá-*, Av. *dasama-*, and Lat. *decimus*). Similarly, 'ninth' was **neun₃-o-* originally, realized phonetically as **neunmo-*, a form only preserved in Latin (*nōnus*, from **noyenos*).

The forms **septmmo-* and **dekmmo-* were early on reanalyzed as **septm₃-mo-* and **dek_m-mo-*, and the apparent suffix **-mo-* that they seemed to contain then spread to neighboring ordinals in some branches, as in Ved. *aṣṭamá-* and OCS *osmŭ* 'eighth'. But in some branches it was the suffix **-to-* that spread in this way, as in Toch. A *ṣāptānt* 'seventh', Lith. *deviñtas* 'ninth', and Eng. *tithe* 'a tenth part' (tenth is a more recent creation).

7.21. Some of the suffixes used for the ordinals, especially **-mo-* and **-to-* (and **-isto-* in the Germanic words for 'first'), are identical to superlative suffixes. This overlap is seen in some of the higher numerals too, as in Vedic *ṣata-tamá-* '100th', using the same suffix *-tama-* as is used for superlatives. See also §6.77.

Adverbs

7.22. To judge by the evidence of the oldest daughter languages, PIE did not possess suffixes whose sole purpose was to change an adjective into an adverb (like Eng. *-ly*); rather, it used case-forms of nouns and adjectives in adverbial function. This

method of forming adverbs is widely productive in all the daughter languages; since most such adverbs are relatively recent creations, few if any are of PIE date. One likely candidate is the neuter nominative-accusative singular of the adjective for 'great' used to mean 'greatly', **meǵh₂* (Hitt. *mēk*, Ved. *māhi*, Gk. *méga*, ON *mjǫk*). Adverbs formed from other case-forms can be exemplified by Lat. *rursus* 'back(wards)' (nominative), *meritō* 'deservedly' (ablative), and *quā* 'how' (old instrumental).

7.23. Multiplicative adverbs for 'twice' and 'thrice' are securely reconstructible as **duis* and **tris*: Ved. *duis*, *tris*, Av. *biš θriš*, Gr. *dis tris*, Lat. *bis ter* (< earlier *terr* < **ters* < **tris*).

Negation

7.24. The two nasals *n-* and *m-* begin certain adverbs expressing negation in the daughter languages. We can reconstruct a sentence negator **ne* 'not', reflected most directly in Germanic (e.g. OE *ne*), Balto-Slavic (e.g. OCS *ne* and Lith. *nė*), Ved. *ná*, and Lat. *ne-* (as in *ne-fās* 'not right, wrong'). The zero-grade of this word was used as a prefix, **n-* 'un-' (Hitt. *am-miyant-* 'not grown, young', Skt. and Gk. *a(n)-*, Lat. *in-*, Eng. *un-*). This prefix is called the privative prefix. Another negator, probably **mē* (as in Ved. *mā*, Gk. *mē*, and, with different initial consonant, Lat. *nē* and Hitt. *lē*), was used with unaugmented (injunctive) verb forms in negative commands.

But many of the negative morphemes in the daughter languages are of more recent vintage and come from words or expressions that were not originally negative in force. People have a tendency to add additional emphasis to negatives (as in Eng. *not at all*, *not a whit*), and often the emphatic addition becomes reinterpreted as the actual negative if the phrase becomes stereotyped. A classic example of this comes from French: in the modern language, verbs are negated with a preceding *ne* (historically the true negator, from Latin *nec*) and a following *pas*. The latter comes from the Latin noun (accus.) *passum* 'step' and was originally used to strengthen the negative force of *ne* with verbs of motion: 'not a step' = 'not at all'. But at some point the phrase *ne ... pas* lost its literal meaning and the word *pas* was felt to convey negativity all on its own, whence its use in phrases like *pas mal* 'not bad'.

The older IE languages have many examples of this development as well. Probably the most famous is Greek *ou* or *ouk(i)* 'not'. The late American Indo-Europeanist Warren Cowgill incisively etymologized this as coming from a pre-Greek phrase **ne oju kʷid*, essentially 'not on your life, not ever': **ne*, the historically "real" negator + **oju* (< PIE **h₂oju*) 'life, age', used adverbially and added to strengthen the negative + **kʷid*, an emphatic particle. Ultimately the negative **ne* lost its salience and was dropped, and **oju kʷid* became **ojukid* (by the *boukólos* rule, §3.39), whence (by regular sound changes) *oukí* and the shorter forms *ouk* and *ou*.

Prepositions and Postpositions

7.25. A good number of words in PIE can be reconstructed that had both adverbial and postpositional function. As postpositions they followed their object, rather than preceding it as prepositions do; but in most of the daughter branches these postpositions eventually became prepositions. Only in Anatolian, Indo-Iranian, Sabellic, and vestigially in Latin and Greek is the old postpositional placement still seen. See further §8.8.

The following are some roots ancestral to many familiar prepositions; as always, descendant forms are given selectively. Unglossed descendant forms mean the same thing as the PIE form. Since prepositions usually develop idiomatic usages, the original sense of some of these roots is not certain.

- **apo* 'from': Ved. *ápa* 'away, forth', Gk. *apó* 'from', Lat. *ab* 'from', Eng. *of, off*
- **en* 'in': Gk. *en*, Lat. *in*, Arm. *i*, Welsh *yn*, Eng. *in*, OPruss. *en*, OCS *vŭ(n-)*
- **en-ter* 'within, inside' (derivative of **en*): Ved. *antár* 'between', Lat. *inter* 'between, among', OIr. *eter* 'between', OHG *untar* 'between, among', Alb. *ndër* 'between, in'
- **epi* or **opi* 'upon, by': Ved. *ápi* 'by, on', Gk. *epí* 'on', Lat. *ob* 'on, over against', Arm. *ew* 'and'
- **kata* or **kṛta* 'down': Hitt. *katta* 'with', Gk. *katá* 'down'
- **kom* 'with': Lat. *cum* 'with', OIr. *co* 'with', probably Gmc. prefix **ga-* 'with'
- **ndher*, **ndhos* 'under': Ved. *adhás*, Av. *adairi*, Lat. *infr-ā*, Arm. *and*, Eng. *under*
- **ni* 'down, under': Ved. *ní*, Arm. *ní*; suffixed in Eng. *ne-ther*, OCS *ni-zŭ* 'down'
- **per(i)* 'around, through': Ved. *pári* 'around, forth', Gk. *perí* 'around', Lat. *per* 'through', OPruss. *per*, Alb. *për*
- **prō* 'forth': Hitt. *parā*, Ved. *prá*, Gk. *pró*, Lat. *prō*, OIr. *ro-* (prefix and intensive particle), Goth. *fra-*, OPruss. *pra, pro* 'through'
- **uper* 'above': Ved. *upári*, Gk. *hupér*, Lat. *s-uper*, OIr. *for*, Arm. (i) *ver* 'up', Eng. *over*
- **upo* 'below': Ved. *úpa* 'up to', Gk. *hupó* 'below', Lat. *s-ub*, OIr. *fo*

Conjunctions and Interjections

7.26. Conjunctions tend not to be stable vocabulary items; new ones get created afresh out of other words. So, for example, English *and* developed out of an adverb or preposition cognate with Gk. *antí* '(over) against', and *but* out of a preposition or adverb meaning 'outside, except'. Similarly, we saw in the list above that the Armenian conjunction *ew* 'and' is derived from an adverb meaning 'upon' (**epi*). A few conjunctions have been reconstructed for PIE, however, that are widely represented in the daughter languages. At least two of them were postpositive particles, being placed directly after the word (or first word of the phrase or clause) that was being conjoined or disjoined: the conjunctive particle **k^we* 'and' (OHitt. *-ku*, Skt. and Av. *ca*, Gk. *te*, Lat. *-que*, Goth. *-h*) and the disjunctive particle **uē* 'or' (Skt. *vā*, Gk. (w) *ē*, Lat. *-ue*). (A famous example of the postpositive usage is the opening of Vergil's *Aeneid* in Latin, *arma uirumque canō*, 'Arms and the man I sing'.) An adversative conjunction **h₂(e)u*, probably 'but, however', is found in Ved. *u*, Gk. *au*, and Lat. *au(-tem)*. For conjunctions introducing embedded (subordinate) clauses, the languages (and surely PIE itself) seem to have resorted to case-forms of various pronouns. Thus from the relative/interrogative stem **k^wo-* we have a neuter accusative in Hitt. *kuit* 'because' and Lat. *quod* 'because', and a masculine accusative in Archaic Lat. *quom* 'when, since' (Classical *cum*).

The particle **k^we* was also used as a generalizing particle, as evidenced by e.g. Hitt. *kuiški* 'someone', Ved. *kás ca* 'everyone, someone' (vs. *kás* 'who?'), Lat. *quisque* 'each, every' (vs. *quis* 'who?'), and Goth. *hazuh* (< **k^wos-u-k^we*) 'every' (vs. *hwas* 'who?').

7.27. The languages also possess a variety of sentence-connecting particles, often translatable as 'and', two of which can be confidently reconstructed for PIE: **nu* (preserved in Hittite and Tocharian B *nu*, in Vedic and Greek *nú*, and as *no-* in Old Irish and *nŭ* in Old Church Slavonic) and postpositive **de* (Gk. *dé*).

Particles

7.28. A number of unstressed words in the daughter languages used as sentence particles seem to be related to each other, although their functions – both in the daughter languages and in PIE – are often difficult to ascertain. A form **kem* or **kṃ* is found in Hitt. *-kan*, Gk. *ke(n)*, and Ved. *kām*; it was perhaps a modal particle. An emphasizing particle **tar* is found in Luv. *-tar* and Homeric Gk. *tar* (usually written incorrectly as *t'ár* in editions of Homer); it appears to have been especially used with interrogatives (Luv. *kuiš-tar* = Homeric Gk. *tís tar* 'who [indeed]?'). Another emphasizing particle **ge* appears frequently after personal pronouns in Greek, as in *egôge* 'I (for one), as for myself, I . . .', accusative *emége*, and may be the source of the velars found in the accusative of the first and second singular pronouns in certain languages: Hitt. *ammuk*, *tuk* 'me, thee', Gmc. **mik*, **þik* (becoming e.g. German *mich*, *dich*). Another particle **ghe* or **gho* is found in Ved. *(g)ha*, *ghā*, and OCS *že*. Lithuanian *ai*, a particle used when asking questions, and the Greek particle *ára* or *rhá*, may be related to each other. Finally, mention may be made of an emphatic particle **ke* (related to the pronominal stem **k̑i-*, §7.9): Hitt. *ki-nun* 'now', with the same elements reversed in Latin *nun-c* 'now'; Lat. *hī-c* 'this'; and Gk. *ekeĩnos* 'that' (**e-ke-eno-*; compare **ke-eno-* in ON *hánn*, *hann* 'he').

Interjections

7.29. It is perhaps indicative of some of the less pleasant aspects of life in the older IE societies that the one PIE interjection we can confidently reconstruct is an expression of woe or agony, **yai*, preserved in Old Hitt. *uwai*, Lat. *uae*, Welsh *gwae*, Goth. *wai*, Eng. *woe*, Latv. *vai*, and, with further suffix for additional pathos, Av. *vaii-ōi*.

Many IE languages have a form like Eng. *oh!*, for which an ancestor **ō!* has been proposed. The laryngealistic rewriting of this as **eh₃!* requires no comment.

For Further Reading

Few book-length treatments of the words covered in this chapter exist except for the personal pronouns and the numerals. The reconstruction of the personal pronouns has been the subject of a fine recent book, Katz 2003, which advances a host of new ideas. An earlier book-length study is Schmidt 1978. For the numerals, one should consult Gvozdanović 1992, which is a sundry collection of articles by different scholars on the numeral systems of each branch; the articles are not all of the same quality, but some are very valuable. A single-authored book-length treatment of issues pertaining to IE numerals is Szemerényi 1960, which should be treated with caution but has much interesting material. Warren Cowgill's etymology of Greek *ou* is in Cowgill 1960.

Exercises

- 1 What are the differences between the nominal and pronominal declensions?
- 2 Identify the following pronominal forms or stems. Be as specific as possible.

a *so *seh ₂ *tod	f *nos	k *moi
b *me-	g *ki-	l *ns-
c *(H)io-	h *ju-	m *k'o-
d *uos	i *el-	n *uei
e *sue-	j *te	o *tu-

- 3 Provide the PIE reconstructions for the following pronominal forms or stems:
 - a 'other'
 - b 'I'
 - c 2nd person sing. dat. enclitic pronoun
 - d 1st person sing. genit. (not enclitic)
 - e 1st person sing. accus. enclitic pronoun
- 4 Discuss the etymological difference(s) between Ved. *eká-*, Greek *heĩs*, and Eng. *one*.
- 5 Provide the PIE reconstructions for the following numerals:

a six	c three	e nine	g hundred
b ten	d twenty	f four	h five
- 6 What do the prefixal forms *tri-, *sm̥-, *ŋ-, and *h₁su- 'good' (§6.86) have in common?
- 7 Provide the PIE reconstructions for the following forms:

a in	c not	e and	g below
b or	d thrice	f forth	h greatly
- 8 Describe the formation of adverbs in PIE.
- 9 Comment on the Latin forms *quisquis* and *quisque* and their congeners from elsewhere in the IE family.
- 10 Describe Warren Cowgill's etymology of Greek *ou* 'not'.

8 Proto-Indo-European Syntax

Introduction

8.1. *Syntax* is the set of rules and principles for the combination of words into larger units – phrases, clauses, and sentences. The term also refers to the branch of linguistics that is concerned with syntactic systems as objects of study. In the comparative historical linguistics of the IE languages, we enter the field of *diachronic syntax*, the study of how syntactic systems change over time. This field and the allied enterprise of syntactic reconstruction raise certain methodological and practical issues that do not arise in diachronic phonology and morphology. We will briefly outline these issues here and then discuss them further below in §§8.12ff.

Morphemes and words are memorized and entered into the mental lexicon during language acquisition, but phrases, clauses, and sentences are not (except for fixed idioms). In the course of ordinary conversation, sentences are produced that have never been uttered before; clearly, a memorized body of sentences, no matter how extensive, would be of little use on its own in producing novel ones. This is how we know that there is a separate syntactic component of the grammar that encodes the procedures for generating phrases, clauses, and so forth. The rules constituting the syntax are deduced from the sentences that are heard during language acquisition in childhood.

8.2. Most contemporary syntactic theories are either wholly or partly the product of several decades of research spearheaded since the 1950s by Noam Chomsky, and constitute the field known as *generative grammar* or *generative syntax*. Though diverging from each other in many respects, all generative syntactic theories assume that the production of sentences is not done simply by stringing words together. First of all, sentence structure is not flat; rather, sentences are hierarchically organized into discrete units called *constituents* that have a particular internal structure and combine with one another in particular ways. In the sentence *The ugly duckling swam with the swans*, the noun phrase *the swans* is a constituent that is part of a larger constituent, the prepositional phrase *with the swans*, which in turn is part of the verb phrase (predicate) *swam with the swans*, which, finally, is combined with the subject noun phrase, *the ugly duckling*, to form a complete sentence.

8.3. According to linguistic theory, generating a sentence proceeds in several stages; in simplified terms, they are as follows. First, the sentence's constituents and subconstituents are arranged in a basic, *underlying* order; this order constitutes the sentence's *deep structure*. The underlying order may then be modified by the application of *movement* rules acting upon particular elements of the sentence.

After these modifications, the relevant phonological and morphological rules are applied, with the final result being the *surface structure*, i.e., what is actually uttered.

Syntacticians try to deduce deep structures and the syntactic rules that act upon them by analyzing surface structures and gathering judgments from native speakers on the grammaticality of particular sentences. Since no native speakers of dead languages are around anymore, the whole enterprise of syntactic analysis of the older IE languages might appear to be doomed from the start. To be sure, the absence of native speakers will always impose certain limits on our knowledge; but that is no reason not to try to figure out as much as we can about the underlying machinery of those languages, utilizing what we know from the syntax of languages alive today. It so happens that many of the preserved texts are long and plentiful enough to allow quite a bit of solid theoretical syntactic analysis, and our analytical methods for doing this are only improving.

8.4. Textual genre is often raised as a problematic issue in these analyses. Many of the oldest preserved literatures, such as the Rig Veda, are works of art whose style and mode of expression can be phenomenally inventive, both lexically and grammatically. However, this problem is sometimes overstated; it is incorrect to suppose – as many have – that poetic texts leave grammar by the wayside, and that poets were able to take “licenses” willy-nilly. The language of poetry is just as strictly rule-governed as ordinary speech: though certain constructions only occur in poetry (leading some scholars to speak of a *poetic grammar*), they are still possibilities afforded by the grammar of the language.

Some texts, especially the Gothic Bible and many Classical Armenian works, are translations that adhere closely to the word order and grammatical constructions of the original language. It is important when analyzing these texts not to make generalizations based on passages that are simply slavish imitations of another language. But since the match in word orders or grammatical constructions is never absolutely exact, and the differences can reflect particular syntactic features in the language of the translation, such texts are still of value in IE syntactic studies.

8.5. Only a few refined and theoretically sophisticated analyses of the syntax of older IE languages are available at present. This fact, coupled with our still limited (but growing) understanding of how syntax changes, prevents us from reconstructing PIE syntax in the same way that we can reconstruct PIE phonology and morphology. However, various syntactic patterns have long been known to be common to most or all the older IE languages, and were likely characteristic of PIE too; some of these will be outlined and exemplified below. Although much of nineteenth- and early twentieth-century syntactic studies of the IE languages focused on analyzing and cataloguing the functions of grammatical categories, such as the cases and moods, some figures of that time, especially Jacob Wackernagel and Berthold Delbrück, also made important discoveries concerning the behavior of constituents and word order. Though not formulated in the same terms as nowadays, their discoveries, such as Wackernagel’s Law (§§8.22ff. below), have been extremely influential.

We will now proceed to survey more specific features usually thought to be part of the PIE syntactic system. Some abbreviations of textual references are not explained here, but can be found in the relevant later chapters.

Syntax of the Phrase

Noun phrases

8.6. Adjectival modifiers of nouns had to agree with their nouns in gender, number, and case. The languages differ as to whether modifiers usually precede or follow their nouns, and the order was surely flexible in PIE also; discourse and pragmatic factors dictated the relative ordering of words. Adjectives typically precede their nouns in Hittite, Vedic, and Germanic, and follow their nouns in Italic and Celtic; but in all these languages there are exceptions. In Celtic and Italic, for example, certain adjectives precede their nouns, while in Hittite, Vedic, and Italic, any noun can precede a modifier when one or the other element was being emphasized.

8.7. Common to all the older languages was the ability of nouns and their modifiers to be separated by intervening elements, yielding what are called discontinuous or distracted noun phrases (a construction called *hyperbaton* in Greek and Latin grammar): Cuneiform Luv. alati awienta Wiluṣati 'they came from steep Wilusa' (KBo 4.11:46), Gk. ándra moi énepe Moṓsa polútropōn 'tell me, Muse, of the resourceful man' (Odyssey 1.1), Lat. magnā cum laude 'with great praise', Arm. zyovs hatanem kenac 'I cut off the hope of life', Old Irish Márta for shúaig saithiu 'on the swarm of the host of March' (Félire Oengusso, March 31). Distraction of other types of phrases was common as well. The technical details of distraction are not well understood; in some cases, it is the result of part of a phrase being moved to a position of emphasis or contrast. Other cases are of a different nature; we will discuss one below in §8.10.

Prepositional and postpositional phrases

8.8. The elements traditionally classified as prepositions were most likely simply independent adverbs in PIE, a status they still largely have in Anatolian, Indo-Iranian, and the oldest Greek. From the comparative evidence it is not entirely clear whether these forms were prepositions, or rather occurred after their objects as postpositions; probably both patterns were current, though many researchers assume postpositional usage to be older. Anatolian and Vedic have almost exclusively postpositions and not prepositions, as in Hitt. šubḫi šēr 'on the roof' and Ved. jānām ānu 'among men'. Avestan and Sabellic have a mixture of prepositions and postpositions. In the other older IE languages, prepositions are the rule, although some (e.g. Old Persian, Greek, and Latin) have vestiges of postpositional use, like Gk. toútōn perí 'about these things' and Lat. mēcum 'with me'.

Preverbs

8.9. Such adverbs, when used to modify the content of verbs, are called *preverbs*. In PIE, preverbs were still independent words, as reflected in Anatolian and older Indo-Iranian. In the examples below, the preverbs and verbs are underlined. (In the Hittite examples in this chapter, an equals sign is a modern orthographic convention for separating enclitic particles from preceding material. The other spelling conventions for transliterating Hittite, such as the use of capitalization, will be introduced in the next chapter and may be ignored for our present purposes.)

Old Hittite *š=aš šarā URU-ya pait¹*
 “and he went up to the city”

Vedic *abhi yó mahinā dīvaṃ mitrō babbhūva saprāthāḥ²*
 “Mitra the renowned who is superior to heaven by his greatness”

Old Avestan *frō mā sāstū vahištā³*
 “let him teach me the best things”

¹ KBo 22.2 rev. B 17' (Zalpa tale). ² Rig Veda 3.59.7. ³ Yasna 45.6.

In the other branches, the preverbs became genuine prefixes attached to the verb. But archaic texts in Greek, Latin, and Old Irish show remnants of the older situation: Gk. *edētios eks éron hénto* ‘they put aside desire for food’ (*Iliad* 1.469 and elsewhere), Archaic Lat. *ob uos sacro* “I entreat you” (Festus 206 L.; would be *uos obsecro* in Classical Latin), OIr. *ad- cruth caim cichther* “fair form will be seen” (*Serglige Con Culainn* 694). Old Irish and Gothic allow certain clitics to intervene between preverb and verb, although the whole complex forms a single phonological word (the clitics are boldfaced in the following examples): OIr. *aidom-indnastar* ‘that I be brought’ (Milan glosses 39^d 13), Goth. *frab ina ga-u-lva-sehvi* ‘He asked him if he saw anything’ (Mark 8:23). The so-called separable prefixes in Modern German and Dutch are either a direct continuation of the PIE situation, or comparable to it: German *auf-tragen* ‘to assign’, *er trägt auf* ‘he assigns’ = Dutch *op-dragen*, *hij draagt op*.

In languages where the preverb and the verb usually form a single word, such as Classical Greek, the separation of the two is called *tnesis*, literally ‘a cutting’. The term is misleading; it arose from a traditional bias toward Classical Greek, which has been viewed as a standard from which Homeric Greek is “deviant,” when in fact the classical language’s loss of *tnesis* is what departs from the inherited situation.

8.10. *Tnesis* is a subtype of the general phenomenon of distraction discussed in §8.7, and examples of it can provide clear documentation of how not all instances of distraction are syntactically equal, except superficially. In the Avestan clause *apa vā yasāiti* ‘or takes (it) away’ (Yasna 11.5), the preverb *apa* ‘away’ is separated from its verb *yasāiti* ‘takes’ only because the clitic conjunction *vā* ‘or’ phonologically requires a single word to its left and therefore must intervene between the preverb and verb. (We will return to such clitics and their behavior in §8.22.) Syntactically quite distinct is a case like *pairi uši vāraiaθβom* ‘cover (their) ears’ (Yasht 1.27), where the preverb (*pairi*, literally ‘around’) has been fronted to the beginning of the clause for prominence or emphasis (by a process discussed in §8.19).

The *vāyav indrás ca* construction

8.11. It was apparently a rule of PIE grammar that when two vocatives were conjoined, the one preceding the conjunction was put in the nominative rather than the vocative case. Almost all the examples of this come from Vedic, as in the phrase *vāyav indrás ca* “o Indra and Vayu” after which the construction is named. In this example, the god Vāyu’s name is in the vocative but Indra’s is in the nominative, as it precedes the conjunction *ca* ‘and’. The sole example outside Indo-Iranian is from an archaic passage in the *Iliad* (3.277): *Zeū páter . . . Éēliós te* ‘Father Zeus (voc.) . . . and Helios (nomin.)’.

Syntax of the Clause

8.12. The reconstruction of the clausal syntax of PIE, for much of its recent history, has been concerned with establishing the “basic word order” of PIE and its daughters. This research program was born of linguistic typology, a field launched in the 1960s that has sought to determine universal patterns of linguistic structure. Most commonly the term “basic word order” has referred to the statistically most frequently observed word order in a given language (another sense of the term will be discussed further below). In the opinion of many scholars today, this approach (often called the typological approach) has led to an intellectual dead end. As a methodological preliminary, we devote some space here to a critique of it so that its mistakes can be learned from (not uncommonly, those same mistakes continue to be made).

First, word order is not equivalent to syntax, but is rather a byproduct of it. Word orders result from the application of syntactic rules, which arrange sentential elements into different functional positions. Pure statistical tallying up of word orders thus not only ignores actual syntax, but also ignores all the factors that intersect with syntax and cause different syntactic rules to apply under different discourse, pragmatic, and semantic conditions. In comparative linguistics, the typological approach has the additional problem of ignoring a fundamental aspect of the comparative method – namely, the identification of comparable structures in related languages on which to base reconstructions. In the domain of morphemes and words, that means cognate forms; in the domain of phrases and sentences, that means phrases and sentences that (broadly speaking) are used when talking about the same sorts of things in the same ways.

For example, both possible relative orders of verb and object (verb–object, object–verb) in the Homeric epics occur quite frequently. However, if one looks at a semantically restricted and well-defined sample, such as traditional sayings or proverbs, there is much more rigidity: the object almost always precedes the verb. The syntax of proverbs in Homer can then be compared with the syntax of proverbs in Vedic Sanskrit and other related languages, and it turns out that these sorts of sentences agree remarkably closely in their syntax across the older IE languages. Proverbs are a repository of traditional language, and often contain constructions that are rare elsewhere; and it is just those kinds of exceptional cases that are most valuable for the comparative method and for syntactic reconstruction (as we discussed in §4.1 with respect to morphology). In English, one syntactic fossil like *till death do us part* tells us more about the history of English syntax than a thousand sentences of newspaper prose.

8.13. A somewhat more modern definition of the term “basic word order” is a single underlying order in the deep structure from which all surface word orders in a language can be derived. According to many mainstream syntactic theories, the statistically most frequent surface word order is not necessarily the same as the basic underlying order. Word orders that are marked (that is, those that encode additional semantic content such as expressive contrast or the like) are derived by manipulation of basic word orders through syntactic movement; but, depending on the language (and the syntactician), unmarked orders may also differ from underlying orders. A refined analysis of texts in the daughter languages should be able to distinguish pragmatically neutral orders from those that are marked, and allow us to come up with syntactic movement rules that can then be compared with rules

in other daughter languages. The existence of syntactic rules for the same semantic or pragmatic purpose (as for example the fronting of a verb to the beginning of a sentence for emphasis or contrast, discussed below in §8.19) in most or all the daughter languages could then be evidence for such a rule in PIE itself.

8.14. It is almost universally asserted that most of the ancient IE languages were verb-final, and that PIE was as well; more specifically, that they were SOV (Subject–Object–Verb). This assertion is made on the basis of sentences such as the following (the verbs are underlined):

- Hittite *nu=za^{MUS} illuyankaš^{DIM} an tarapta¹*
 “And the serpent overcame the Stormgod.”
- Vedic *maruto ha enam na ajahub²*
 “Indeed the Maruts did not abandon him.”
- Latin *Eumolpus tanquam litterarum studiōsus utique atramentum habet³*
 “Eumolpus, so interested in learning, surely has (some) ink.”
- Runic *ek hlewagastiz holtijaz horna tawido⁴*
 “I, Hlewagastiz of Holt, made (this) horn.”
- Tocharian A *kāsu ñom-klyu tsrašissi šāk kälmentwam sätkatär⁵*
 “Good fame of the strong spreads out in ten directions.”

¹ KBo 3.7 i 11 (Illuyanka myth, §3). ² Aitareya-Brahmaṇa 3.20. ³ Petronius, Satyricon 102.

⁴ Gallehus horn, normalized transliteration (see §15.36). ⁵ Puṇyavanta-jātaka 1 (see §17.32).

As a matter of fact, the claim that many or most of the older IE languages were verb-final has never been clearly verified. Part of the problem with it is arriving at a clear definition of a verb-final language. In the strict sense, a verb-final language is one where the verb always comes at the end of each clause unless other factors intervene. The only well-known older PIE language that meets this criterion is Hittite. No matter what the genre, no matter how stylistically marked the text, the verb in Hittite is always clause-final, with one exception – when it is fronted to the beginning of the clause for emphasis or contrast (by the fronting rule that we will get to shortly). None of the other old IE languages behaves so rigidly (except for Insular Celtic, but that subbranch contains verb-initial languages; see §8.19); there is essentially no position in the clause (on the surface at least) where the verb cannot appear.

It is usually stated that in these languages, the pragmatically neutral order is SOV. This may, in fact, be true, at least of some of them (such as Latin); but with so many word-order permutations possible (and frequent), clearly they cannot be called “verb-final” in the same way as Hittite. There are any number of reasons, according to current theory, why a verb may or may not appear as the last word in its clause. Syntactic rules place the constituents of a clause in certain functional slots; superficially different linear positionings of a constituent can turn out, on closer examination, to be identical underlyingly, and vice versa.

The foregoing paragraphs, it must be repeated, do not represent the majority view, but are also not intended to claim that that view is wrong – just that it needs tighter formulation and convincing demonstration.

8.15. One usually thinks of clauses or sentences as containing a verb by definition, but in the older IE languages, many of their modern descendants, and doubtless in PIE itself, clauses could lack an overt verb. Chief among these are nominal sentences, in which the copula (the linking verb 'be') is understood but not overtly expressed: OHitt. *annaš=šiš* MUŠ-aš (KUB 1.16 ii 20) 'his mother (is) a snake' (recall §8.9 for the use of the equals sign), OPers. *manā pitā Vištāspa* 'my father (is) Vištāspa' (DSf 12–13), Gk. *emoi d' ákhos* 'and to me (there is) pain' (*Iliad* 5.759), Lat. *tū coniūnx* 'you (are) his wife' (Vergil, *Aeneid* 4.113), Toch. A *tsrašiñ waste wraśāši* 'the strong (are) the protection of creatures' (Punyavantajātaka 2).

Subject-verb agreement

8.16. In clauses containing a verb, the subject (when expressed) had to agree with the verb in person and number. An apparent exception to this is the behavior of neuter plural subjects in Anatolian, Old Avestan, and Greek, which in these languages take singular verbs. In the following examples, the subjects are bold and the verbs are underlined: Hitt. *apē=ya uddār* (pl.) QATAMMA *lagāru* (sing.) 'let also those words fall over likewise' (KUB 2.3 iii 21–22), Old Avestan *yā* (pl.) *zī vānuarazōi* (sing.) 'the things which have been perpetrated' (Yasna 29.4), Greek *hōssa te phúlla* (pl.) *kai ánthēa* (pl.) *gígnetai* (sing.) *hórēi* 'as many as the leaves and the flowers that appear in their season' (*Iliad* 2.468). But this phenomenon is likely due ultimately to the ancient status of the neuter plural as a collective; recall its history, discussed in §6.68.

8.17. Because the personal endings of verbs in the older IE languages encode the subject in them, it was not grammatically necessary to use an overt personal pronominal subject in addition. However, subject pronouns nonetheless occur; in the traditional grammars of these languages, overt subject pronouns are usually characterized as emphatic. (Terms like "emphasis" are rather blunt and can get overused, but they are a useful starting point on the way to more refined analyses.) When the Roman historian Sallust wrote (*Bellum Catilinae* 6.1) *sicūtī ego accēpī* 'as I understand it', the overt subject pronoun *ego* 'I' serves to contrast his own understanding with the opinion of others. While such contrast is certainly evident in many instances where a subject pronoun is expressed, it is not clear that overt subject pronouns are limited to contrastive or emphatic use, and it is often difficult or impossible to ascertain just why a writer used one. In nominal sentences, an overt pronominal subject is generally required for clarity, as in Old Persian *adam navama* 'I (am) the ninth' (DB 1.10).

8.18. The syntax of the reflexive possessive adjective **suo-* 'own' (§7.12) deserves a special note. Reflexive adjectives (and pronouns) refer back to the grammatical subject of a sentence. But the possessive **suo-* had broader usage, to judge by the daughter languages: it could refer back not to the grammatical subject, but to newly introduced discourse material or to an older topic that is returned to. As an example of the former, consider Rig Veda 8.2.7: *trāya índrasya sómāḥ sutāsaḥ santu devāsya své kṣāye sutapāvnah* 'Let the three somas be pressed for the god Indra in the soma-drinker's own house' (translation following Brent Vine; soma was an intoxicating sacred drink). Here the grammatical subject is *trāya* . . . *sómāḥ* 'the three somas' and the possessive *své* refers to the soma-drinker, who is newly introduced. Similar behavior can be found in other older IE languages.

Basic movement processes

8.19. When a constituent or subconstituent appears in a position other than the usual one, syntacticians usually treat it as having been placed there by a movement rule. (Actually, many “usual” positionings of elements are also thought to arise by movement processes.) Typically, elements are moved leftward rather than rightward, by a process called **fronting**. The technical details do not concern us, but based on evidence from a variety of languages it is believed that the grammar provides a number of slots or “landing-sites” in the abstract syntactic representation of a sentence, which serve various functions and into which elements can be moved or to which elements can be adjoined. One such landing-site has been posited for the left edge of a sentence into which a constituent or subconstituent can be moved for emphasis or contrast; the process of fronting into this position is called **topicalization**. Some examples of clauses in which the verb has been topicalized follow (the emphasis or contrast does not always come across in translation):

- Hittite *ḫalziššai=wa=ita* DINGIR^{MEŠ}-aš *attaš* ^D*Kumarbiš*¹
 “Kumarbi, the father of the gods, is calling you.”
- Old Av. *sraōtū sāsnaš fšāghiiō suiē taštō*²
 “Let the bondsman (?), fashioned for benefit, hear the teachings.”
- Greek *ménei tò theíon douliái per en phrení*³
 “The divine (power), even when in bondage, stays in the mind.”
- Latin *fui^{mus} Trōes, fui^{it} Ilium*⁴
 “We were (but no longer are) Trojans, Troy was (but no longer is).”
- Armenian *erknēr erkin erknēr erkir*⁵
 “In labor was heaven, in labor was earth.”

¹ KUB 33.122 ii 11. ² Yasna 49.9. ³ Aeschylus, *Agamemnon* 1084. ⁴ Vergil, *Aeneid* 2.325.

⁵ Movsēs Xorenac’i 1.31 (see §16.42).

If verb-initial order generated in this way becomes stereotyped, it can be reanalyzed by learners as the neutral order; and in fact in Insular Celtic, VSO order became the norm for precisely this reason (the perhaps older verb-final order is still the rule in the Continental Celtic language Celtiberian). A similar reanalysis happened in Lycian; see §9.68. Certain verbs, especially existential verbs (e.g., ‘there is’) but also verbs of speaking and imperatives, preferentially occur clause-initially across all the IE languages: Skt. *āsīd rājā nalo nāma* ‘there was a king named Nala’ (*Mahābhārata* 3.53.1), Lat. *est in cōspectū Tenedos nōtissima fāmā īnsula* ‘within sight there is a most famous island, Tenedos’ (Vergil, *Aeneid* 2.21–22), dialectal Old Russian *estī gradū mežu nobomū i zemleju* ‘there is a city between heaven and earth’ (Novgorod birch bark fragment 10.1).

Clause-initial position, as already noted, is a place of prominence not only for verbs, but for any constituent, as the following Hittite examples show: *irma=šmaš=kan dābhun* ‘sickness I have taken away from you’ (KBo 17.1 i 12’), *anmel=ma tarnumar ŪL pāi* ‘but my release he won’t give’ (KBo 32.15 iii 16), *šarā=kan namma ehu* ‘come up again!’ (KUB 33.84+ iv 11’). Topicalization was probably a syntactic process in PIE.

8.20. Another type of fronting that occurred in all known ancient IE daughter languages and is widely known from living languages too is WH-movement. This process has also been posited for the proto-language. WH-movement gets its name from the initial letters of English interrogatives such as *who*, *which*, *what*, which were shown in early generative syntactic work often to originate in deep structure in a later syntactic position than the one they occupy on the surface. Subsequent research has indicated that the syntactic slot to which interrogatives move (now called the *complementizer* position) can be occupied by other kinds of elements too, including relative pronouns and subordinating conjunctions. The complementizer position precedes the rest of the clausal positions proper, which is why in English and many other languages conjunctions such as *before*, *because*, *if*, *when*, and so forth appear first in their clause.

8.21. However, at least according to one analysis, in the older IE languages the complementizer position was preceded by the topicalization position; if the latter was filled by a topicalized element, the complementizer was no longer clause-initial. The following examples contain both a complementizer and a topicalized element; the topicalized element is underlined, the complementizer boldfaced:

- Hittite *ammuaq=a kuit ḫarkun*¹
 "And also (that) which I had"
- Vedic *jātām yād enam apāso ādhārayan*²
 "when the craftsmen held him, just born"
- Old Avestan *naēnaēstārō yaθanā vohunam mahi*³
 "since we are non-scorners of good things"
- Latin *fēstō diē sī quid prodēgeris*⁴
 "if you splurge a bit on a holiday"

¹ Apology of Hattusili III, iv 69. ² Rig Veda 3.2.7. ³ Yasna 35.2. ⁴ Plautus, *Aulularia* 380.

This marks a fundamental difference in syntactic behavior between the older IE languages and a language like English. In some languages, particularly Latin, syntactic movement can distribute numerous elements before the complementizer position, resulting sometimes in surface orders where all the words in a clause except one precede the complementizer (and, in some unusual poetic styles, where the complementizer even comes at the end of the clause). Whether this is due to the presence of multiple landing-sites in front of the complementizer position in this language, or to some other process, is not yet known. In the following conditional clauses from the Archaic Latin of Plautus, the subordinating conjunction *sī* 'if' can be preceded by some or all the other clausal constituents save the verb:

- saluos domum sī redierō* "if I shall have returned home safe"¹
mē quoque unā sī cum illō relinquerēs "if you were to leave me together with him too"²
perfidia et peculātus ex urbe et auāritia sī exulant "if betrayal and embezzlement and greed are exiled from the city"³

¹ *Amphitruo* 584b. ² *Captivi* 446. ³ *Persa* 555.

Wackernagel's Law and the placement of clitics

8.22. Perhaps the most famous feature of the clausal syntax of older IE languages is the tendency of unstressed clitic particles to appear second in their clause after the first stressed element, a phenomenon discovered by Jacob Wackernagel in the late nineteenth century. Observe the clitics (underlined) in the following sentences:

Hittite *kiell=a parnaš ēšhar papratar QATAMMA pattenuddu*¹
 “Of this house too may it likewise drive out the bloodshed (and) uncleanliness.”

Vedic *ā tvā mántrāḥ kaviśastā vahantu*²
 “Let the spells recited by the poets lead you hither.”

Greek *ēmos d’ ērigéneia phánē rhododáktulos Eōs*³
 “but when early-born, rosy-fingered Dawn appeared”

Latin *tū autem in neruō iam iacēbis*⁴
 “But you will soon be lying in custody.”

Gothic *fram-uh þamma sokida Peilatus fraletan ina*⁵
 “And at this Pilate sought to release him.”

¹ KUB 7.41 ii 54–55. ² Rig Veda 10.14.4. ³ *Iliad* 1.477 and passim. ⁴ Plautus, *Curculio* 718. ⁵ John 19:12.

Sometimes, however, one of these clitics appears as the third or fourth word in its clause. Recent research, especially by the American linguist Mark Hale, has shown that Wackernagel's Law actually involves several processes that usually, but not always, conspire to place unstressed particles in second position in the clause. His discoveries have explained the exceptions to a strict formulation of the law. A simplified account is presented here.

8.23. Three types of postpositive clitics (and clitic-positioning rules) can be distinguished. *Word-level clitics* modify or limit (or have scope over, in technical parlance) a single word or constituent, and are placed directly after the word or the first element of the constituent. Such clitics tend to have the function of emphasizing the word to which they are attached, or setting it in some kind of contrastive relief or focus (the clitic is boldfaced): Hitt. *nu=wa=za **apun**=pat eši* ‘occupy only **that** (land)’ (KUB 14.1 obv. 19), Ved. *pracyāvāyanto **acyutā** cid* ‘the ones who move even **unmovable things**’ (Rig Veda 1.85.4). If the word that such a particle modifies is first in its clause, then the particle appears (coincidentally) second in its clause: Ved. ***sthirā** cid ānnā dayate ví jāmbhaiḥ* ‘even **rough** food he cuts apart with his teeth’ (Rig Veda 4.7.10), Lat. ***hoc quoque** maleficium* ‘**this** crime too’ (Cicero, *Pro Roscio Amerino* 117). Such particles, when modifying a phrase, can often come second in the phrase, as in Gk. *én ge taís Thēbais* ‘in all of Thebes **indeed**’ (Sophocles, *Oedipus Tyrannus* 1380). Some clitics, such as the descendants of PIE **k^we* ‘and’, can act as word-level clitics as well as sentence connectors (see the next paragraph).

8.24. *Sentence-connective clitics* conjoin or disjoin clauses or sub-clausal constituents. Examples of these clitics are the various descendants of PIE **k^we* ‘and’ and **uē* ‘or’ (see §7.26). They are attached to the first word of the constituent or clause

being conjoined or disjoined, whether that is a single word (Ved. *ághna índras ca* ‘o Agni and Indra’, Rig Veda 3.25.4), a phrase (Lat. *silua alta louis lúcusue Diānae* ‘the high forest of Jupiter or the grove of Diana’, Vergil, *Aeneid* 3.681), or a clause (Old Avestan *yā xī āgharē yāscā hantī yāscā mazdā buuaintī* ‘indeed (those) who were and who are and who will be, o Mazda’, Yasna 33.10).

8.25. Finally, *sentential clitics* are clitics whose scope is a whole clause or sentence. These include the unstressed personal pronouns as well as a variety of sentential adverbs that serve expressive functions and are often untranslatable into English. They are positioned in various ways. Some are placed after the first stressed word in a sentence and any emphatic or sentence-connective clitics associated with that word, while others (called “special clitics” in the technical literature) are positioned after a particular syntactic structural position in the clause. If the first word in a sentence is a proclitic, that is, an unstressed word that attaches phonologically to a following stressed word, the sentential clitic will of course not come directly after it, as in Gk. *eks hēmēōn gār phāsi kāk’ ēmmenai* ‘for they say that bad things are from us’ (*Odyssey* 1.33), where the proclitic *eks* ‘from’ is not a proper phonological host for the clitic *gār*. Sentential clitics occur not infrequently in strings or chains: Ved. *nā vā u etān mriyase* ‘indeed you do not die thereby’ (Rig Veda 1.162.21), Gk. *ē rhā nú moi ti pūthoio* ‘may you indeed now trust me somewhat’ (*Iliad* 4.93), Hitt. DUMU-ŠU=*ma=wa=šši=za=kan* ‘but his son himself to him . . .’ (all the clitics after *-ma* ‘but’ are sentential; only *-šši* ‘to him’ and *-za* ‘himself’ are translatable).

Subordinate clauses

RELATIVE CLAUSES

8.26. Relative clauses in numerous daughter languages share certain characteristics that are worth remarking on and are probably inherited. In the older IE languages, the relative clause often precedes the main clause (and with it, the antecedent). The relative pronoun or adverb is often paired with a pronominal or adverbial antecedent, yielding what are called correlative structures of the type ‘(the one) who . . . , he . . .’ or ‘in the way which . . . , in that way . . .’. In the examples below, the relative clauses are underlined, and the relative pronoun and antecedent are boldfaced:

Vedic *yéna imā víśvā cyávanā kṛtāni . . . sá janāsa índrah*¹
“(The one) **by whom** all **these things** have been made to shake . . . that, people, (is) Indra.”

Old Av. *at yāng ašātacā vōistā vanhēušcā dāθōng mananhō ərəθβāng mazdā aburā aēibiio parānā āpanāiš kāmam x^ā2*
“(Those) **whom** you know (to be) just in accordance with truth and with good thought (and) worthy, **Ahura Mazda**, fulfill for them (their) desire with profits.” (tr. adapted from Humbach; see ch. 11)

Greek *hós ke theoīs epipeithētai, málā t’ ékluon autoi*³
“**Whoever** obeys the gods, they listen to him as well.”

¹ Rig Veda 2.2.4. ² Yasna 28.10. ³ *Iliad* 1.218.

Very characteristically, if the antecedent is a noun rather than a pronoun, it is placed within the relative clause and in the same case as the relative, sometimes repeated in

the main clause. Thus instead of saying *The gods who gave us riches can take them away*, speakers of these languages would have said literally, *Which gods gave us riches, they/those gods can take them away*:

Hittite *nu=kan kās IM-aš kuēz wappuwaz danza nu zik wappuaš* ^DMAH
 *tuēl ŠU-TIKA dā*¹
 (lit.) “from which riverbank this clay (has been) taken, o genius of
 (that) riverbank, take (it) in your hand”, i.e. “O genius of the
 riverbank from which this clay has been taken . . .”

Archaic Latin *quem agrum eōs uēdere herēdemque sequi licet, is ager uectigal nei*
 *siet*²
 “the field which (lit., which field) they are allowed to sell and pass
 to an heir, that field may not be taxable.”

¹ Ritual of Tunnawi 1.30–32. ² *Sententia Minuciorum*, CIL I² 584.5.

8.27. As can be seen from some of the examples so far quoted, the relative pronoun did not need to be the first member of its clause. In several of the ancient IE languages, the relative could be preceded at least by a topicalized element, just like the subordinating conjunctions discussed above. In Insular Celtic, the relative pronoun became fixed in second position of the clause, able to intervene between a preverb and following verb: Old Irish *ní latt aní ara-rethi* (Würzburg Glosses 6^b 22) “it is not yours that which you assail” (*ara-rethi* < Celtic **are-yo retesi*, where **-yo* is the relative pronoun). This happened prehistorically in Baltic as well: the inherited PIE relative pronominal stem **iō-* became attached to adjectives to make them definite (see next paragraph), and there are a few Old Lithuanian examples of participles beginning with preverbs where the relative pronoun occurs not at the end of the participle, but after the preverb, such as *nu-jam-lūdusam* ‘saddened’ (masc. dative sing. of the past active participle).

8.28. Independently in some daughter languages, the relative pronoun became reanalyzed as a marker of definiteness of an attributive adjective. Already in Old Avestan one comes across constructions that would have been an early stage in this development, such as *tāiš śīiaoθmāiš yāiš vahištāiš* (Yasna 35.4) ‘with the best works’, lit. ‘with the works which (are) best’. The relative pronoun *yāiš* has been attracted into the instrumental plural, the case of its antecedent *śīiaoθmāiš* ‘works’; earlier the construction would have been **tāiš śīiaoθmāiš yā vahištā*, with the relative pronoun in the nominative plural (as subject of its clause) and the predicate adjective *vahištā* also in the nominative plural. Old Persian has examples like *karā haya manā* ‘my army’, lit. ‘army which (is) mine’, with an indeclinable relative *hya*; this developed ultimately into the Modern Persian *ezāfe* construction, as in *ketāb-e naw* ‘the new book’, where the particle *-e* marks definiteness of an attributive adjective and possession (*ketāb-e man* ‘my book’). In Balto-Slavic, definite adjectives have an extra syllable beginning with *-j-*, from the PIE relative pronoun **iō-*: OCS *novyjǫ zavětŭ* ‘the New Testament’, Lith. (accus. sing.) *naująjį vardą* ‘the new name’. The Anatolian language Carian has recently been analyzed as having a similar construction, in phrases like *šwinś upe arieś-χi ted* (M30) ‘stele of Šwin, father of Arie’, where *arieś-χi ted* literally means ‘of Arie who (is) the father’ (with *-χi* continuing the other PIE relative pronominal stem, **k^wi-*).

Negation

8.29. As detailed in the previous chapter, for ordinary negation PIE had an adverb **ne* and a privative prefix **n-* ‘un-’. In some cases, the privative prefix was used where one might expect the conjunction. Comparative evidence suggests that certain classes of words were preferentially negated not with the adverb but with the privative prefix; among these words were participles and verbal adjectives. Greek and Latin, for example, ordinarily use their negative adverbs when negating participles, but some fixed archaic constructions point to an earlier time when the privative prefix was used instead, as Homeric Gk. *a-ékontos emeio* ‘with me being unwilling, against my will’, Lat. *mē in-sciente* ‘with me not knowing’, *in-uītus* ‘unwilling’ (later replaced by *nōn uolēns* ‘not willing’), *im-prūdēns* ‘not knowing beforehand’ (later *nōn prouidēns*). Compare also Av. *an-usant-* ‘not wanting’, Goth. *un-agands* ‘not fearing’.

In clauses, the positioning of negation is complex. If the negation has scope over a single word or constituent, it usually directly precedes that constituent. Sentential negation typically directly precedes the verb, as in English. But it could also be moved toward the front of the sentence for emphasis.

Absolute constructions

8.30. An action, state, or event could be syntactically backgrounded using a construction called an absolute. Typically the absolute consisted of a noun modified by a participle – semantically equivalent to a subject plus verb – in an oblique case. Thus Latin has ablative absolutes (*hīs rēbus gestīs* lit. ‘these things having been done’, i.e. ‘after these things were done’ or ‘because these things were done’), Greek has genitive absolutes (Homeric *aékontos emeio* ‘with me being unwilling’), Vedic Sanskrit has locative absolutes (*ucchāntyām uśási* ‘with dawn shining forth’), and Gothic and Old Church Slavonic have dative absolutes (Goth. *imma rodjandin* and OCS *jemu glagoljōšĉemu* ‘with him speaking, while he is/was speaking’). PIE surely had such constructions too, although which case or cases were used is debated. Opinions also differ as to how many of these are native in the languages in which they are attested.

Phrase and Sentence Prosody and the Interaction of Syntax and Phonology

8.31. An area that has received considerable attention in theoretical linguistics since the 1980s is the interaction of syntax and phonology, the so-called “phonology–syntax interface.” Speech sounds in ordinary speech are grouped into hierarchically arranged levels of organization; this organization is termed the *prosody* (not to be confused with the rules of poetic meter, also called prosody). Each level of the prosody is called a *prosodic domain*. Syllables constitute the smallest prosodic domain; they combine to form words, which combine with clitics to form clitic groups; clitic groups and words, or words and words, form phonological phrases, which themselves combine to form intonational phrases. Each of these units is a level to which particular phonological rules are applied. For example, in clitic groups in Vedic Sanskrit, a *t*

will become *n* before a nasal, as in *prá tán me voco* 'proclaim that to me' (underlyingly *tát me*; Rig Veda 7.86.4). This same rule does not apply within the smaller prosodic domain of the word, as in *ātmá* 'breath'. Within any prosodic domain larger than the word, a phonological rule will affect the *junctures* between words; such rules are usually called *sandhi rules*. Common types of phonological processes that can happen at word junctures include assimilation (as in Ved. *tán me* above, and as in the pronunciation of Eng. *I hit you* as [aj hitʃu]) and resyllabification (as when *an ice man* sounds identical to *a nice man* in fluent speech).

8.32. The construction of prosodic domains above the level of the word is influenced by syntactic configuration. Under the usual assumptions about the mapping of syntactic structure onto the prosody, words belonging to the same constituent that start out as contiguous in the deep structure and stay contiguous throughout the derivation will tend to be grouped together as a single phonological unit, whereas words that only become contiguous through certain kinds of movement sometimes do not. In other words, syntactic movement can block the application of a particular sandhi rule, or can prevent a word from turning into a clitic. For example, in Greek, clitics normally receive no accent, but if two or more occur in a string, all but the last one get accented, as in *ei mē tis me theōn* 'if no one of the gods me...' (Odyssey 4.364). However, in a sequence like *douliai per en phreni* 'even in bondage in the mind' (Aeschylus, *Agamemnon* 1084, quoted above in §8.19), there are two clitics in a row but the first is not accented. The reason is that *per* 'even' emphasizes *douliai* 'in bondage' and is phonologically attached to it, while *en* 'in' is a preposition that governs *phreni* 'the mind' and is proclitic to that word. The two resultant clitic groups [*douliai per*] and [*en phreni*] form two separate prosodic groups with what is called a prosodic boundary between them. (A prosodic boundary, incidentally, is not generally audible as a pause or other break.) We conclude that the rule placing an accent on the first of two successive clitics applies only if the two clitics belong to the same clitic group.

8.33. The writing systems of most languages do not consistently reflect the output of sandhi rules or related prosodic processes, but they do not entirely ignore them, either. In Hittite, chains of clitics (see §9.13) are written without word-breaks between the clitics, indicating that they all formed a clitic group that acted as a single phonological word. In medieval Irish manuscripts, prepositions are written together with following nouns as one word, indicating a similar prosodic fact for prepositional phrases. Greek and Latin inscriptions frequently show evidence of sandhi rules and prosodic groupings that are not reflected in the more rigid orthography of the standard literary tradition. For example, interpuncts (punctuation marks separating words) in Greek inscriptions are ordinarily not placed between a definite article and a following noun, or between a preposition and a following object. Latin inscriptions preserve spellings such as *quot per* 'because through (...)' where standard orthography demanded *quod per*; the less literary inscriber simply wrote as he spoke – with devoicing of the final *-d* before the following voiceless *p*, a sandhi rule about which the standard orthography provides no information. Sometimes authors who are linguistically more sensitive than most devise a phonetic spelling system that reflects external sandhi and related phenomena, such as the Old High German author Notker and the early Middle English author Orm (see §15.64). Most exceptional of all is the Sanskrit writing system (called Devanagari; see §10.57), which encodes sandhi phenomena in great detail; the received text of the Rig Veda has been of immense value in shedding light on the interaction of syntax and prosody in older IE languages.

8.34. One widely observed effect of the syntax on phonology involves the “heaviness” or “weight” of constituents. Noun phrases consisting of a bare noun, for example, are much more likely to enter into certain kinds of clitic groups than are noun phrases where the noun is modified by another element (called branching noun phrases). In punctuated Greek inscriptions, as we saw above, interpuncts do not ordinarily separate a definite article from a following noun; but an interpunct is present if the article is followed by a branching noun phrase, indicating a stronger prosodic break between the two. In Homer, we see a difference in the behavior of different types of prepositional phrases vis-à-vis the positioning of the sentence-connecting conjunction *dé*, an enclitic that normally occurs second in its clause. If the clause begins with a prepositional phrase consisting simply of a preposition plus bare noun, the clitic will follow the whole phrase (e.g. *eks pántōn dé* ‘and of all’, *Iliad* 4.96), whereas if the clause begins with a more complex phrase consisting of a preposition followed by an adjective-noun phrase, the clitic will come in between the preposition and the rest of the phrase (e.g. *dià dè khrōa kalón* ‘and into the fair flesh’, *Iliad* 5.858). This has been interpreted to show that *eks pántōn* is prosodically cohesive enough to function as a single word for the purposes of clitic placement, while *dià khrōa kalón* is not. A similar phenomenon is found in the Old High German of Notker Labeo: a definite article is written without an accent when preceding a simple noun phrase (e.g. *taz héiza fiur* ‘the hot fire’, J28), indicating clisis and destressing of the article, but is written with an accent when preceding more complex noun phrases (e.g. *die uuilsalda állero búrgo* ‘the fortune of all cities’, J19).

8.35. Metrics – the study of the rules and behavior of poetic meters – can be a valuable source of information about prosodic phonology. In certain Greek and Roman meters, for example, there is a rule that a sequence of two light syllables in particular verse-positions must belong to the same word. The rule, though, has an interesting exception: a word-break between the two syllables is allowed when one of them belongs to a proclitic (as in the sequence *ut opiniōne* ‘that in [his] opinion’, Plautus, *Miles Gloriosus* 1238). This means that the prosodic group consisting of proclitic plus word was tighter than that consisting of two full-content words – tight enough to behave, for the purposes of the poetic meter, as though there were no word-division. In the Old English epic poem *Beowulf*, whose poetry – like all older Germanic heroic poetry – is structured around stressed words alliterating with one another, weakly stressed and unstressed words like prepositions, pronouns, and certain adverbs do not participate in alliteration. Additionally, some verbs do not participate in it either, thus patterning with weakly stressed elements – a fact that is in agreement with the general propensity of verbs in IE to be more weakly stressed than nouns (as we saw in §5.59). Much of the poetry in the older IE languages remains to be exploited for similar information; the classical languages and Vedic have been most intensively studied in this regard but have themselves by no means yielded all their secrets.

8.36. It is reasonable to suppose that phonological and morphophonological processes such as prosodic domain formation and cliticization were conditioned in PIE by structural factors like those described in this section. There is as yet no theory of diachronic prosody – how prosodic systems change historically in the course of a language’s development – and accounting for the phonology-syntax interface is the subject of ongoing debate in contemporary generative linguistics. As our understanding of some of these questions improves, and as the ancient IE texts

are more closely scrutinized for the prosodic information they can provide, we will hopefully be able to determine which rules can be projected back onto PIE itself.

For Further Reading

Most grammars of Indo-European languages and introductory books on Indo-European give short shrift to syntax or neglect it altogether. The fundamental nineteenth- and early twentieth-century works, which are in German, mostly treat the usage of morphosyntactic categories. These include Wackernagel 1920 and Delbrück 1893–1900, a monumental work in three volumes. Wackernagel's Law was proposed in Wackernagel 1892. The modern age of IE syntax was ushered in by Calvert Watkins with his article on the Old Irish verbal complex (Watkins 1963); see also Watkins 1976, which lays out the problems with purely typological approaches to IE syntax and word order. Generative linguistic approaches have informed more recent studies such as Hale 1987 and Garrett 1990, both important and influential works. On the “phonology–syntax interface” in general, see the articles in Inkelas and Zec 1990.

For Review

Know the meaning or significance of the following:

constituent	<i>vāyav indraś ca</i>	fronting	sentence-connective
underlying	construction	topicalization	clitic
movement	preverb	WH-movement	sentential clitic
deep structure	tnesis	Wackernagel's Law	prosodic domain
surface structure	nominal sentence	word-level clitic	sandhi rule

Exercises

- Briefly describe the syntactic behavior or significance of the following forms or constructions mentioned in the chapter:

a Lat. <i>magnā cum laude</i>	g PIE * <i>kʷe</i>
b Lat. <i>mēcum</i>	h Hitt. DUMU-ŠU=ma=wa=šši=za=kan
c Gk. <i>Zeū páter . . . Éállós tē</i>	i Lat. <i>quem agrum . . . is ager . . .</i>
d Hitt. <i>annaš=šiš MUŠ-aš</i>	j Gk. <i>a-ékontos emeio</i>
e Skt. <i>āśid rājā nalo nāma</i>	k Ved. <i>prā tán me voco</i> (vs. <i>ātmā</i>)
f Hitt. <i>-pat</i>	l OHG <i>dīe uuilsakda állero búrgō</i>

- Determine the syntactic rule that is responsible for the position of the underlined forms in each of the following sentences from the Rig Veda:

- divó no vṛṣṭīm maruto raridhvam* (5.83.6) “Give us the rain of heaven, o Maruts”
- kvà idānīm sūriah* (1.35.7) “Where is the sun now?”
- yī suparṇō antāriṣṣāni akhyad* (1.35.7) “The bird has looked out over the sky's regions” (*akhyad* = ‘saw’)

d rátham kó nír avartayat (10.135.5) "Who set the chariot rolling down?"
(kó = 'who')

- 3 The sentence in 2b above literally translates "Where now (the) sun". What kind of sentence is this?
- 4 In §8.8, the Greek postpositional phrase *toutón péri* was mentioned. Where do you suppose the accent was in PIE, based on this evidence and on the discussion in §8.8?

9 Anatolian

Introduction

9.1. The Asian part of Turkey, called **Anatolia in classical times**, and parts of northern Syria are the home of the earliest attested Indo-European languages, together called **Anatolian**. Anatolian was not established as a branch of Indo-European until the twentieth century. Two letters found at Tell el-Amarna, Egypt, and written in Mesopotamian cuneiform script were the first documents to come to light that contained **Hittite**, now the most famous Anatolian language. Known as the Arzawa letters, they were published in 1902; the editor, Jorgen Knudtzon, identified the language as Indo-European, but this view was generally greeted with disbelief. In 1906, the first large-scale Hittite finds were made, in the central Turkish village of Boğazköy (now Boğazkale), roughly 90 miles east of Ankara. Excavations in that year uncovered about 10,000 cuneiform tablets and tablet fragments dating from the mid- to late second millennium BC; these finds still represent the bulk of our material in the language. It turned out that Boğazköy is on the site of **Hattusas**, the capital of the Hittite Kingdom (see below); the tablets were found in the remains of a royal archive. During the First World War, the Czech philologist **Bedřich Hrozný** deciphered the language and proved that it was Indo-European.

9.2. Since then other Anatolian languages, none as well preserved as Hittite, have been identified and have gradually been giving up their secrets: **Palaic** and **Cuneiform Luvian**, both also from the second millennium BC and written in the same cuneiform script as Hittite; **Hieroglyphic Luvian**, with remains from both the second and the first millennia BC and written in a native hieroglyphic script; and the first-millennium languages **Lycian** and **Lydian**, written in a Greek-derived alphabet. (Lycian was already known by the late nineteenth century, but its recognition as Anatolian did not come until later.) Three other languages of southern Anatolia – **Carian**, **Pisidian**, and **Sidetic** – surely belong to the family as well, but due to their fragmentary nature they are very poorly known. As we have come to understand the non-Hittite languages better, their importance for Anatolian and IE studies has increased tremendously. In the past, the focus of attention was the relationship between Hittite and IE, without concern for the intermediate stage of Common Anatolian; but we are now in a better position to reconstruct that stage.

9.3. Anatolian is one of the growth industries of IE studies. Almost since the time of its discovery, it has engendered vigorous debate about the structure of PIE and the subgrouping of the family. Our understanding of these matters is being continually revised because new inscriptions are discovered on a regular basis, and our interpreta-



Map 9.1 Anatolia

tions of already-known texts and linguistic forms are constantly improving. Any historical account of the Anatolian languages is therefore somewhat provisional.

Anatolian, Indo-European, and “Indo-Hittite”

9.4. The main difficulty posed by Anatolian for Indo-Europeanists is the fact that its structure is quite different from that of PIE as traditionally reconstructed. One would ordinarily expect that the oldest known languages of the family should resemble the proto-language the most closely; and it is true that Anatolian does preserve a number of important archaisms, notably consonantal reflexes of at least one (and probably two) of the three laryngeals and productive classes of neuter *r/n*-stems. But most striking are the forms and categories that it does *not* have. Absent are such apparently bedrock IE formations as simple thematic verbs, the aorist, the perfect, the subjunctive, the optative, the dual, the **-tó*-verbal adjective, and the comparative in **-ios-*. Additionally, some inherited grammatical forms function differently from their congeners in the rest of IE: endings identifiable with the IE perfect are used to form a class of presents (the so-called *hi*-conjugation; §9.12); and the **-nt*-participle, which has active voice in the other IE languages, is passive in Anatolian.

One interpretation of these facts is that the forms missing from Anatolian were simply lost, and that the traditional reconstruction of PIE is perfectly valid. But evidence has been growing that Anatolian split off at a time when the development of some of these categories (such as the *s*-aorist) was only nascent. Under this view, PIE went through some subsequent development before achieving its “classic” look; the “missing data” of Anatolian are then partly attributable to loss, and partly to their not having existed yet. An early version of this theory was propounded by the American Indo-Europeanist Edgar Sturtevant, who thought Anatolian was a sister of reconstructed PIE and that both were derived from a language he called “Indo-Hittite” (a term that later fell out of favor).

From PIE to Common Anatolian

Phonology

9.5. Anatolian is justly famous for being the only branch to preserve consonantal reflexes of the laryngeals. All scholars agree that $*h_2$ has a reflex in Hittite, Luvian, and Palaic as the sound transcribed h , and in Lycian variously as χ (x), k , and q . It is also very probable that $*h_3$ was preserved as h in Hittite, but not all have accepted this hypothesis.

9.6. According to recent assessments of Luvian, Anatolian is also the only branch of IE that preserved the original three-way distinction among the velars.

9.7. As for the other consonants, the voiced aspirated stops lost their aspiration and merged with the plain voiced stops. The liquids and nasals generally stayed intact. Interestingly, no native Anatolian words began with r (except in clearly secondary cases), which has led some scholars to believe that PIE had no r -initial words either. Under this view, PIE words reconstructed with $*r$ - should be reconstructed as $*h_1r$ -. The matter is still debated.

9.8. Anatolian inherited all the PIE vowels, but in every language except Lycian $*o$ and $*a$ fell together as a . Long \bar{o} and \bar{a} also fell together, as \bar{a} .

Morphology

9.9. Anatolian nouns were richly inflected in the singular. All the PIE cases were represented, as well as a directive (or allative) case in $*-a$, which indicated place to which; the directive was probably a case in PIE also, although most of the evidence for it comes from Anatolian (recall §6.12). An ending $*-ti$ is reconstructed for the Common Anatolian ablative and instrumental, both singular and plural (e.g. Luv. *Wilušati* 'from Wilusa'). In the plural, separate endings are reconstructible only for the nominative, accusative, genitive, and dative in Common Anatolian, and there is no trace of the other oblique plurals with suffixes beginning with $*-m$ - or $*-bh$ - (§6.17). Impressionistically, the plural has a more primitive look than elsewhere in IE. An important innovation was a new "ergative" ending $*-enti$ for neuters, together with the creation of a split-ergative syntax (see below); this ending was originally the instrumental of neuter n -stems.

The archaic r/m -stems (§6.31) are a large and productive class of nouns in Anatolian, reduced to a mere handful of forms in the rest of IE; for example, Luv. *kuršawar*, stem *kuršaw(a)n*- 'island' (from $*kuršai$ - 'cut', i.e. 'land cut off by water').

9.10. Anatolian preserves the distinction between distributive and collective plural in animate nouns (§6.68); thus Old Hittite *alpeš* 'clouds' vs. collective *alpa* 'group of clouds'. No direct reflex of the dual has been found. Whether the feminine exists in Anatolian in any way is debated (recall §6.69).

9.11. Among the striking features of the pronominal system is the use of clitic subject pronouns; elsewhere in IE, clitic pronouns are only found in oblique cases. The Anatolian subject pronouns are limited to use with a few classes of intransitive verbs. Among other features worthy of note in the pronominal system are a demonstrative stem $*obo$ - (< PIE $*obho$ -) meaning 'that' (> Hitt. *apā*-, Pal. *apa*-, Lyc. *ebe*-) and the odd u -vocalism of the 1st person pronoun $*ug$ 'I' and $*emu$ 'me' as opposed to PIE $*eġ$ - and $*eme$ (Hitt. nomin. *ūk*, accus. *ammuk*; cp. Hieroglyphic Luv., Lyc., and Lyd. accus. *anu*). This u ultimately comes from the influence of $*tu$ - 'thee'; cp. §7.1.

VERBS

9.12. Anatolian verbs only inflected in two tenses, present and preterite, and in two voices, active and mediopassive. Thus there are far fewer verbal inflections in Anatolian than in the rest of IE. Among the types of present stems, simple thematic verbs are exceedingly rare or non-existent, but root athematic presents are very common. While the aorist as a category is not found, a few present stems appear to correspond to aorist stems in other languages, such as Lyc. *tadi* 'puts' (cp. Ved. aorist (*ádhat*). There was also an imperative, with 3rd person forms ending in **(t)u* (e.g. Hitt. *pau* 'let him give', *paiddu* 'let him go'; compare Ved. Skt. *étu* 'let him go').

But the morphological peculiarity for which the Anatolian verb is best known is the existence of two separate verbal conjugations, the *mi*-conjugation and the *hi*-conjugation (named after their respective Hittite 1st sing. present active endings). The *-mi* conjugation is unproblematic, continuing the familiar endings **-mi* **-si* **-ti* etc. in the present and **-m* **-s* **-t* etc. in the preterite. The *hi*-conjugation endings, however, are directly equatable with those of the IE perfect. How the *hi*-conjugation relates to the perfect is a major unsolved problem and bears on the very foundations of the IE verbal system. Under an older view, Anatolian turned the perfect into a present tense for one (diverse) group of verbs. Under a newer, still controversial view, the *hi*-conjugation continues a class of PIE presents and aorists that was allied with the perfect in complex ways; see §9.33 below. There are also many other theories.

Syntax

9.13. The Anatolian languages are syntactically quite similar to one another. Characteristic of Anatolian sentence structure is the clause-initial clitic chain: all clitic sentential conjunctions, clitic sentential adverbs, and clitic pronouns were attached to the first word of the clause in a particular fixed order, with the whole chain written as a single word. Thus in Hittite one can begin a sentence *kinun=ma=wa=tu=za* (the equals signs are used to indicate the clitic boundaries in Anatolian scholarship), where *kinun* means 'now', *-ma* is a clitic adversative conjunction meaning 'but', *-wa* is a quotative particle indicating that the sentence is quoted speech, *-tu* is the 2nd sing. accus. pronoun, and *-za* is a reflexive particle. Clauses are generally verb-final, although the verb or any other element in a clause could be fronted to the beginning of a clause for emphasis or topicalization. Also characteristic of Anatolian clauses was the almost universal use of a so-called sentence connective to begin any clause that was not discourse-initial; this conjunction served as an anchor for the clitic chain (in the absence of a fronted element filling that slot) and also connected each sentence to the previous one in continuous narrative. It roughly corresponded to 'and' or 'and then' in English, but is frequently untranslatable. The most familiar example is Hitt. *nu*, cognate with Eng. *now* (but not so translated).

9.14. As noted above, an innovation of Common Anatolian morphosyntax was the creation of a so-called ergative case in **-enti*. This was a case taken by neuter nouns when they functioned as the subjects of transitive verbs, and comes historically from the ablative-instrumental case of neuter *n*-stems. Its development can be understood if we assume that pre-Anatolian originally did not allow inanimate nouns to be the subject of transitive verbs (a phenomenon found in other languages too). Thus instead of saying *The fire burned the house*, one would have said, (It) *burned the house with fire* or the like, with the word for 'fire' in the ablative-instrumental

case. The ablative-instrumental in such constructions was reinterpreted at some point as a different subject case, which we now call the ergative.

Hittite

9.15. Because of the extensive linguistic remains, far more is known about the Hittites than about the other Anatolian peoples; but little can be said about their origins. That they or their ancestors did not originally inhabit Anatolia is certain (in spite of the controversy mentioned in §2.71). Their designation for their country, *Ḫatti*, is taken over from the indigenous non-Indo-European Hattic people whom they encountered early and assimilated. The influence of Hattic culture was quite strong; preserved among the hundreds of ritual texts at Boğazköy are many of Hattic provenance. The Hittites, along with other Anatolian speakers, presumably came from the north, and were already settled in central Anatolia by about 1900 BC. Around this time Assyrian merchants were setting up colonies in the area, and in Assyrian documents from Kanesh (modern Kültepe, about 100 miles south-southeast of Boğazköy) the first Hittite words appear, as loanwords – proper names and a few lexical borrowings, two of which are of some cultural significance: *išpa(t)ta(l)lu* ‘nightwatch’ (Hitt. **išpantalla-*, from *išpant-* ‘night’) and *išḫiu(l)lu* ‘contract’ (Hitt. *išḫul*). The former is interesting in light of the fact that Anittas, the first Hittite king of whom we have records (see below), took enemy cities by night; the latter is revealing in light of the importance of the contract and other mutually binding arrangements in ancient IE societies (see §§2.11ff.).

Cultural and political history

9.16. At the dawn of Hittite history, the central part of Anatolia (the area known as Cappadocia in classical times) was divided into numerous small principalities. Hittite history is ushered in with the “Anittas text,” the oldest surviving Hittite document, dating from the sixteenth century BC or earlier. It contains an account of how one Anittas and his father, Pithanas, king of a place called Kussara, vied against the other principalities for power. They were eventually victorious and set up their capital at Nesa (also called Kanesh). From the name Nesa comes the adverb *nešumnili* ‘in Hittite’.

9.17. The connection between Anittas and the kings of the subsequent Old Kingdom is not known; he is not claimed by any of them as a forebear. The history of the Old Kingdom apparently starts with Hattusilis I, who moved the capital to Hattusas and also began a southward territorial expansion into what is now northern Syria. His son Mursilis I pressed into Mesopotamia and even sacked Babylon. The Old Kingdom culminated with the reign of the powerful king Telipinus in the fifteenth century BC.

During the early years of the **Middle Kingdom**, the state prospered at first but then was beset by invaders from every direction; records even speak of Hattusas getting sacked and burned. A strong king came to the rescue, Suppiluliumas I, who ascended the throne around 1380 BC. With his reign the **New Kingdom** or Hittite Empire began – the most powerful period in the history of the Hittites. For nearly two centuries the Hittite Empire was one of the most influential states in the ancient Near East. Eventually, though, it fell victim to invasions from the north, ending with the destruction of Hattusas around 1200 BC. While this did not spell the end of

Hittite civilization, it did terminate the production of Hittite texts, since Hattusas had been the center of the Hittite scribal chancelleries.

9.18. Various principalities of the former empire remained intact until near the end of the eighth century BC and called themselves Hittites, most prominently a state in northern Syria with its capital at Karkamish (Carchemish) on the Euphrates. Their chancellery language was not Hittite, however, but Hieroglyphic Luvian. After the Assyrians conquered Syria in the late eighth century, we never again hear of anyone called Hittite.

The linguistic stages of Hittite

9.19. Hittite is divided into three periods that correspond to the three political periods outlined above: **Old Hittite**, the language of the Old Kingdom (c. 1570–1450 BC); **Middle Hittite** (c. 1450–1380 BC); and **Neo-Hittite** (c. 1350–1200 BC), beginning with the death of Suppiluliumas I (documents from his reign are linguistically still Middle Hittite). About four centuries separate the oldest Old Hittite from the latest Neo-Hittite – about as much time as separates present-day English from the language of Shakespeare.

Except for one bronze tablet, all surviving Hittite texts are inscribed on clay. The dating of Hittite texts and the description of Old, Middle, and Neo-Hittite have been hampered by the fact that texts were often recopied and modernized. Many tablets inscribed during the New Kingdom contain Neo-Hittite versions of Old and Middle Hittite originals and present a mixture of archaic and modernized forms. Greater accuracy in dating has been achieved largely through advances in *paleography*, the study of sign-shapes and their evolution. We now know in considerable detail the relative chronology of different cuneiform sign-shapes and of various external features of the tablets. The earliest style (or *ductus*) of Hittite cuneiform is called **Old Script**, followed by **Middle Script** and **New Script**.

9.20. The Hittite corpus spans a wide range of genres. Most plentiful are ritual texts, containing instructions (some quite lengthy) on how to conduct the rituals and festivals of the Hittite state religion. Of great importance to historians of the Ancient Near East are the numerous treaties, decrees, annals of kings, official letters, and a sizeable law code. We also have many myths, omens, and medical texts.

Copies of the cuneiform tablets have been published in several collections since the early 1900s; published tablets are identified by an abbreviation of the collection, the volume number, and the tablet number. Thus the catalogue number of the text sample in §9.44 below, KBo 17.1 (or KBo XVII 1), means the first tablet in the 17th volume of the collection *Keilschrifttexte aus Boghazköy* (*Cuneiform Texts from Boghazköy*). Reference may then be made to column and line numbers: KBo 17.1 iii 12 refers to line 12 in column 3. All the Hittite texts were organized into a thematic (genre-based) catalogue by Emmanuel Laroche called CTH (*Catalogue des textes hittites*; Laroche 1971); it is now standard practice to give the CTH number for any text being discussed.

Cuneiform writing and its transcription

9.21. Before proceeding with our historical grammatical sketch, we will introduce the nature of, and problems with, the writing system and the transcription of it. Hittite was written in the Mesopotamian cuneiform script. This script was probably

invented by the Sumerians; Sumerian is at any rate the first language recorded in it that we can read, beginning at the end of the fourth millennium BC. Over the ensuing millennium and a half it was adopted by many other cultures of the Ancient Near East. The designation *cuneiform* comes from Latin *cuneus* 'wedge', after the wedge-shaped marks made by a reed stylus pressing into soft clay. The clay tablet was the writing medium *par excellence* in a region where both clay and the sunshine needed to harden it were abundant.

Cuneiform writing was originally designed for keeping mercantile records and consisted of symbols for numbers and pictorial depictions of trade items. But it eventually spread into other spheres of life, necessitating the representation of abstract concepts for which pictures could not be drawn. As a result, by about 2800 BC the script had developed into a partly phonetic system in which individual signs could be read as *ideograms* (also called *logograms*, representations of whole words or concepts) or as *syllabograms* (representing individual syllables). For example, the sign for the Sumerian word for 'mouth', pronounced *ka*, could be used logographically to represent the concept 'mouth' or syllabographically to represent the syllable *ka* in some other (longer) word. In essence, what the Sumerians developed was a rebus system, similar to English word-games where a picture of an eye can be used to represent the concept 'eye', the pronoun 'I', or the syllable [aj] in some longer word. Over time the system increased greatly in complexity, with some syllables represented by any of numerous signs, and with many signs having multiple values.

9.22. From the Sumerians the script spread by c. 2400 BC to the Akkadians, whose language was Semitic (Sumerian is not certainly related to any known language family). The Akkadians took over the whole system, and continued to use many of the signs in their ideographic Sumerian values as a kind of shorthand; these are called *Sumerograms*. However, since Akkadians read Sumerograms out loud, it was often convenient to the reader to indicate what the inflections of the underlying Akkadian words were. This was done by adding a sign or two indicating the Akkadian inflectional ending. This extra information is known as a *phonetic complement*. (The use of phonetic complements, incidentally, is also characteristic of Japanese writing; and for a parallel from English, compare the addition of letters in the ordinal numerals *2nd*, *3rd*, *4th*, etc.)

One version of Akkadian cuneiform, known as Old Babylonian, was adopted by the Hittites sometime before 1600 BC. They took the above practices one step further: they used both Sumerian and Akkadian words (the latter called *Akkadograms*) as a shorthand, and sometimes attached Hittite phonetic complements to them. Thus in a Hittite text, theoretically any given sign could stand for a Sumerian word, a syllable of a Sumerian word, a syllable of an Akkadian word or phonetic complement, or a syllable of a Hittite word or phonetic complement.

9.23. In transliterating Anatolian cuneiform texts, the three languages Sumerian, Akkadian, and Hittite (or Luvian or Palaic) are conventionally differentiated by contrastive use of capitalization and italics. These and other conventions are best explained using an illustrative example (KBo 3.4 ii 41f.):

na-aš-ma-at-ta URU.KÙ.BABBAR-ša-aš ZAG-aš ku-iš BE-LU ma-ni-in-ku-wa-an nu
ERÍN^{MES} ANŠE.KUR.RA^{MES} a-pé-e-da-ni ú-e-ek-ti

"or whichever border-guard of Hattusas is near you, (if) you ask for infantry and chariot-fighters from him"

Hyphens are used to separate the transcriptions of individual cuneiform signs: thus the first word, *na-aš-ma-at-ta*, is written with five signs. (As we will discuss below in §§9.30 and 31, individual cuneiform signs often represented sequences of two speech sounds.) Signs used to write Hittite are transcribed in lower case (such as the first word); Sumerograms are in roman upper case (as ZAG); and Akkadograms are in italic upper case (as *BE-LU*). When a Sumerian word consists of more than one sign, periods are used instead of hyphens to mark the sign boundaries (as ANŠE.KUR.RA). Hittite phonetic complements are joined to a preceding Sumerogram or Akkadogram with a hyphen: ZAG-*aš* consists of the Sumerogram ZAG ‘border’ plus the Hittite phonetic complement *-aš* indicating genitive singular (the underlying Hittite word is *irḫaš*), and ^{URU}KÜ.BABBAR-*ša-aš* represents *Ḫattušaš* ‘of Hattusas’, where the last syllable of the Hittite word (*-ša-aš*) is added to the Sumerogram as a phonetic complement.

A superscript sign preceding a word, as in this last example (^{URU}), is called a *determiner* (or *determinative*) and indicates the general semantic class to which the word belongs. In this example, ^{URU} (literally ‘city’) indicates the name of a city. The superscript signs that follow a word are usually Sumerian grammatical endings (or Akkadian, if they are italicized); the one occurring in our example (^{MEŠ}) is a Sumerian plural marker.

9.24. The system of transcribing sign by sign shown above, called *narrow transcription*, is often clumsy (although it is used for all editions of Hittite texts); it can be normalized by “erasing” the hyphens and redundant vowels (*broad transcription*). Thus the word *ma-ni-in-ku-wa-an* above can be normalized as *maninkuwan*. In a word like *ḫa-a-ra-aš* ‘eagle’ or *ma-a-an* ‘when, if’, the extra vowel sign for *a* must be taken seriously, for this was how the Hittites wrote long (and possibly also stressed) vowels; these words are normalized as *ḫāraš* and *mān*. (The technical term for the use of an extra vowel sign is *scriptio plena*, ‘full script’). Double consonants are written double in both kinds of transcription, as the difference between single and double consonants is significant (see §9.31 below): *šu-up-pi-iš* = *šuppiš* ‘pure’. In all the forms cited below, broad transcription will be used, as is standard.

The letters used in transcription have their Latin values except *z*, which represents *ts* (as in German). The letters *š* and *ḫ* probably represent *s* and the velar fricative [x] (as in German *Bach*). Diacritics and subscript numerals are used to differentiate homophonous cuneiform signs (e.g. *gú* [or *gu*₂], *gù* [or *gu*₃], *gu*₄, etc.)

Phonology

9.25. Only some of the numerous sound changes undergone by Hittite will be outlined here. In the stop system, the palatal and plain velar stops merged into plain velars as in centum languages; compare *kāš* ‘this’ (< **kos*) with *lukke-* ‘to kindle’ (**loukeje-*), both spelled with *k*. The labiovelars preserved their labial element and are usually written with the sign *ku*: *kuiš* ‘who’ < **k^wis*, *kuenzi* ‘kills’ < **g^whenti*.

9.26. A characteristic consonant change is the development of **t* to *ts* (written *z*) before an *i*. The 3rd sing. and pl. primary personal endings **-ti* and **-nti* thus become *-zi* and *-nzi* in Hittite, as in *ēšzi* ‘is’ and *ašanzi* ‘are’ (cp. Ved. *āsti*, *sānti*). Also characteristic of Hittite is the dissimilatory change of **n* to *m* when next to *u*. Thus the 1st pl. primary verb ending *-weni* appears as *-m(m)eni* in verbs like *arnummeni* ‘we stir, move’.

9.27. Hittite preserves the second and probably also the third laryngeal as *h*, as in *paḫḫur* 'fire' < **peh₂ur* and *ḫaraš* 'eagle' < **h₃er*-.

9.28. The syllabic resonants developed an *a* in front of them, as in *palḫi*- 'broad' < **p_lh₂-i*-, *kard*- 'heart' < **k₁rd*-, and *anzās* 'us' < **nos*-.

9.29. The Common Anatolian vowels underwent complex changes in Hittite, of which we may mention the following. Short **e* became *a* in several environments, especially before sonorants, although the conditioning factors have not been fully worked out. Examples include *anda* 'in, into' < PIE **endo*; *tarku*- 'to dance' < **terk^w*- ('twist' in PIE; cp. the name of the English dance!); and *daššu*- 'massive' < **dens*-. Short **e* also generally became *a* when unaccented: the 1st and 2nd pl. endings *-weni* and *-teni* appear in Old Hittite as *-wani* and *-tani* when the verbal root is stressed. Short accented *ā*: *wātar* 'water' < **uód₁r*, *kunānt*- 'killed' < **g^whnónt*- (with *o*-grade of the participial ending as in Greek *-ont*-).

HITTITE PHONOLOGY AND THE CUNEIFORM SYLLABARY

9.30. In their phonetic values, cuneiform signs constituted a syllabary, not an alphabet: signs stood for syllables and not individual sounds (except vowels). Signs represented single vowels (V), sequences of consonant + vowel (CV), vowel + consonant (VC), or consonant + vowel + consonant (CVC). Not all theoretically possible combinations are found.

Such a spelling system has limitations. Since no cuneiform signs stand for consonantal sounds alone, a syllabary cannot unambiguously represent word-initial or word-final consonant clusters, or word-internal clusters of more than two consonants. From comparative evidence and from spelling fluctuations in Hittite texts, it is certain that Hittite had such clusters; but in many cases we are uncertain about the pronunciation.

For example, Hittite words that go back to PIE forms beginning with the cluster **sp*- are typically written beginning *išp*-. But the verb meaning 'I pour a libation', which goes back to the root **spend*- (cp. Gk. *spéndō* 'I pour a libation'), is normally written beginning *šip*- (*šipānti*). To further complicate matters, its derivatives are written with *išp*-, such as *išpantuzzi* 'libation vessel'. In either or both the *išp*- and *šip*- spellings, the *i* might be an "empty vowel" that was not pronounced; but there is no way of knowing for sure, given the limitations of the syllabary and the fact that the "sprouting" of a vowel before or within such a cluster has happened in many other languages.

9.31. The Akkadian syllabary distinguished voiced from voiceless consonants in its CV signs (e.g., *pa* and *ba*), but the Hittites did not avail themselves of this distinction in any systematic way. Crucially, there is no correlation between the sign used and whether the relevant sound comes from a voiced or voiceless consonant in PIE. For example, *a-ta-an-zi* and *a-da-an-zi* 'they eat' are both attested spellings of a word that etymologically contains *d* (from **h₁ed*- 'eat').

What the Hittites did instead was make a systematic distinction between consonants written single between vowels (V-CV) and consonants written double (VC-CV). Consonants written single correspond to PIE voiced stops, while those written double correspond to voiceless stops. Thus the single *t* or *d* in *a-ta-an-zi* and *a-da-an-zi*, which comes from **d*, contrasts with the double *-tt*- or *-dd*- in a word like *ar-nu-ut-tu* or *ar-nu-ud-du* 'let him bring', which comes from **t*. This fact is known as Sturtevant's Law, after the American Indo-Europeanist Edgar Sturtevant. What these spellings represent phonetically is still being debated.

Morphology

VERBS

9.32. As in Common Anatolian, Hittite verbs were conjugated only in the indicative mood, in the present and preterite tenses, and in the active and middle (mediopassive) voices. There was also an active and middle imperative for all persons. Mention has already been made of the *mi*- and *hi*-conjugations; the endings of both are alike in the plural and middle, and often a verb belonging to one conjugation acquired endings of the other conjugation by analogy, especially in later texts. Sample active paradigms are given below for the *mi*-verb *ēpmi* 'I take, seize' and the *hi*-verb *ārhi* 'I reach, arrive at'. To exemplify the middle, the paradigm of *arḫaḫari* 'I stand' is given; variant forms of the 3rd singular, without *-t-*, are illustrated with relevant forms of *kišḫaḫari* 'I become'. Bracketed forms are not actually attested for these verbs:

		Active		Middle
		<i>mi</i> -conjugation	<i>hi</i> -conjugation	
Present	sg. 1	<i>ēpmi</i>	<i>ārhi</i>	<i>arḫari</i> , ¹ <i>arḫaḫari</i>
	2	<i>ēpši</i>	<i>ārti</i>	<i>artati</i>
	3	<i>ēpzi</i>	<i>ari, āri</i>	<i>arta(ri), kiša(ri)</i>
	pl. 1	<i>appuemi</i>	<i>arweni</i>	<i>arwašta</i>
	2	<i>apteni</i>	<i>arteni</i>	[<i>arduma</i>]
	3	<i>appanzi</i>	<i>aranzi</i>	<i>aranta, arantari</i>
Preterite	sg. 1	<i>ēppun</i>	<i>ārḫun</i>	<i>arḫati</i> , ¹ <i>arḫaḫat</i>
	2	<i>ēpta</i>	[<i>ārta</i>]	<i>artat(i)</i>
	3	<i>ēpta</i>	<i>āraš</i>	<i>artat, kišat(i)</i>
	pl. 1	<i>ēppuen</i>	<i>erwen</i>	<i>arwaštat</i>
	2	<i>ēpten</i>	[<i>erten</i>]	[<i>ardumat</i>]
	3	<i>ēppir</i>	<i>erir</i>	<i>arantat, arantati</i>

¹ Old Hittite.

To the active forms above may be added the imperatives, which in the 2nd singular are the base stem (*ēp*) and in the 2nd plural end in *-ten* (*ēpten, arten*); in the 3rd person, a *-u* appears in place of the *-i* of the indicative (sing. *ēpdu, aru*; pl. *appandu*).

9.33. There is considerable variation in ablaut in root athematic verbs of both conjugations. For example, alongside the forms *appuemi apteni* above are also found *ēppuemi* and *ēpteni*. The 3rd singular present tense of ablauting *mi*-verbs often has *e*-vocalism that contrasts with either *a*-vocalism or no vowel in the plural: compare the singular/plural pairs *ēšzi* 'is' ~ *ašanzi* 'they are', *ēkuzi* 'drinks' ~ *akuwanzi* 'they drink', *kuenzi* 'kills' ~ *kunanzi* 'they kill'. Some ablauting *hi*-verbs have the opposite pattern: *šakki* 'knows' ~ *šekkanzi* 'they know'. Others have *a*-vocalism throughout, sometimes with a contrast of single versus double consonant: *āki* 'dies' ~ *akkanzi* 'they die'.

An *a* ~ *e* alternation in Anatolian likely continues an **o* ~ **e* alternation in PIE. This is reminiscent of the fact that several dozen verbal roots show an interchange of *o*- and *e*-grade in cognate present-tense forms in other daughter languages: Goth. *graban* 'to dig' (< **ghrobh-*) vs. OCS *po-grebo* 'I dig' (< **ghrebh-*); OIr. *melid* 'grinds'

(< **melh*₂-) vs. Goth. *malan* 'to grind' (< **molh*₂-); Gk. *a(w)éksomai* 'I grow' (< **h*₂*ueks*-) vs. Goth. *wahsan* 'to grow' (= Eng. *wax*; < **h*₂*uoks*-). To explain these verbs and the ablaut of the *hi*-conjugation, it has been suggested by the American Indo-Europeanist Jay Jasanoff that PIE had a class of presents, the "*h*₂*e*-conjugation," with *o*-grade in the singular and *e*-grade in the dual and plural that originally inflected with endings like the perfect. This idea has not found universal acceptance but has been gaining adherents, and the research it has spawned has been shedding light on the relationships among the *hi*-conjugation, the perfect, and the middle.

9.34. Besides root athematic verbs, there are several types of derived stems. Two kinds of nasal presents are found, both conjugating as *mi*-verbs, those with infix *-nin-* and those with suffix *-nu-*; both form causatives, such as *harmink-* 'destroy' (< *hark-* 'perish') and *armu-* 'bring' (< *ar-* 'reach' above). The PIE factitive suffix **-eh*₂- (§5.34) is faithfully continued in Hittite as *-ahh-*, as in *idālawahh-* 'do bad, do evil' (< *idālu-* 'evil'); the factitives conjugate first as *hi*-verbs and later also as *mi*-verbs. Among thematic verbal suffixes, one that is particularly widespread is *-(i)ške-* (< PIE **-ske-*), which could be freely attached to any verb stem to form what is standardly called an iterative. The term is misleading, for – as alluded to in §5.31 – the suffix was used to indicate not only repeated action, but also background action, action lasting over a long time (durative), and action undertaken by several subjects or applied to several objects (distributive). Thus *uw-anzi* 'they look' ~ *ušk-andu* 'let them (all) look', *memai* 'says' ~ *memiškezzi* 'says (many things)'.

9.35. Hittite verbs also had an infinitive in *-anna* (e.g. *adanna* 'to eat') or *-uwanzi* (e.g. *iyawanzi* 'to bring'), used much like the English infinitive; a supine in *-uwan*, used only with the verb *dāi-* 'place' in a construction meaning 'begin to...' (e.g. *memiškiwan dāi* 'he begins to speak'); and the *nt*-participle, which (unlike elsewhere in IE) is a past passive participle when added to transitive verbs (e.g. *kunānt-* '[having been] killed' from *kuen-* 'kill') but active when added to intransitives (e.g. *pānt-* '[having] gone' from *pāi-* 'go'). It is functionally equivalent to the PIE **tó*-verbal adjective, which is not found in Anatolian.

9.36. That Hittite inherited mobile accentual paradigms is clear from the shift of long vowels in such pairs as Old Hitt. *dāhhe* 'I take' ~ *tumēni* 'we take' (**dóh*₃-*h*₂*e* vs. **dh*₃-*yéni*) and *ēšzi* 'is' ~ *ašānt-* 'having been' (**h*₁*és*-*ti* **h*₁*s*-*ónt*-).

NOUNS

9.37. A representative selection of Hittite noun and adjective declensions is given below. Paradigms are given for *antuhšaš* 'man', *halkiš* 'grain', *haššuš* 'king' (attested only in the singular), *idālu-* 'evil', *hūmant-* 'every, all', and *uttar* 'word, speech'. All are common-gender except *uttar*.

	<i>a</i> -stem	<i>i</i> -stem	<i>u</i> -stem noun	<i>u</i> -stem adj.	<i>nt</i> -stem	<i>r/n</i> -stem
singular						
N	<i>antuhšaš</i>	<i>halkiš</i>	<i>haššuš</i>	<i>idāluš</i>	<i>hūmanza</i>	<i>uttar</i>
V			<i>haššue</i>			
A	<i>antuhšan</i>	<i>halkin</i>	<i>haššun</i>	<i>idālun</i>	<i>hūmandan</i>	<i>uttar</i>
G	<i>antuhšaš</i>	<i>halkiaš</i>	<i>haššuwaš</i>	<i>idālawas</i>	<i>hūmandas</i>	<i>uddanas</i>
D-L	<i>antuhši</i>		<i>haššui</i>	<i>idālawi</i>	<i>hūmantı</i>	<i>uddanı</i>
Ab	<i>antuhšaz</i>	<i>halkiyaz</i>	<i>haššuwarz</i>	<i>idālawaz</i>	<i>hūmandaz</i>	<i>uddanasz</i>
I		<i>halkit</i>		<i>idalawit</i>		<i>uddanit</i>

	<i>a</i> -stem	<i>i</i> -stem	<i>u</i> -stem adj.	<i>nt</i> -stem	<i>r/n</i> -stem
plural					
N	<i>antuhšeš</i>	<i>ḫalkiš</i>	<i>idālawēš</i>	<i>ḫūmanteš</i>	<i>uddār</i>
A	<i>antuhšuš</i>	<i>ḫalkiuš</i>	<i>idālamuš</i>	<i>ḫūmanduš</i>	<i>uddār</i>
G	<i>antuhšaš</i>		<i>idālawāš</i>	<i>ḫūmandāš</i>	<i>uddanaš</i>
D-L	<i>antuhšaš</i>		<i>idālawāš</i>	<i>ḫūmandāš</i>	<i>uddanaš</i>
Ab			<i>idālawaz</i>		

9.38. As in Common Anatolian, Hittite nouns distinguished more cases in the singular than in the plural. Over the course of the language's attested history, certain case distinctions in the singular disappeared as well. The ablative took over the functions of the instrumental, and the directive disappeared, its function taken over by the dative-locative.

The genitive plural in Old Hittite ended in *-an* (< **-ōm*), as in *šūman* 'of the gods', but this got replaced by the *-aš* of the dative-locative, as in the paradigms above. The accusative plural ending *-uš* appears to be the regular outcome of PIE **-ns*; the *u*-stem accus. pl. *-amuš* is dissimilated from **-awuš* (from PIE **-eu_h-ns*) by the sound change noted above in §9.23. Not represented in the particular paradigms above is the Old Hittite directive which ended in *-a* (as in *arunā* 'to the sea'). Also not represented is the ergative, which ended in *-anza*, as in *paḫḫuenanza* 'fire'.

9.39. Hittite, as well as the other Anatolian languages, has a much more robust class of *r/n*-stems than the other IE languages. Aside from such basic inherited nouns as *paḫḫur* genit. *paḫḫuenāš* 'fire', *ēšḫar išḫanaš* 'blood', and *wātar wetenaš* 'water', derivatives using the suffixes *-ātar* and *-ēššar* were freely formed, e.g. *paḫrātar paḫrannaš* 'defilement' and *uppeššar/uppešnaš* 'sending, message'.

9.40. Remains of mobile accentual paradigms are preserved in such pairs as *wātar* (sing.) 'water' ~ *widār* (collective pl.) 'waters' (< **uód_hr* ~ **uedōr*) and *tēkan* (nominative sing.) 'earth' ~ *taknī* (dative sing.) (< **dhēghōm* ~ **dhēghmēi*).

Syntax

9.41. Like the other Anatolian languages, Hittite clauses (except those at the start of a discourse or a major section in a text) typically begin with a sentence connective to which one or more clitics could be attached in a clitic chain. The most common sentence connective was *nu*; Old Hittite also had *ta* and *šu*. The initial position in the sentence, i.e. the place held by the sentence connective, could also be filled with another element in the sentence that was topicalized or fronted for emphasis or contrast; and such an element could also host the clitic chain. Various examples can be seen in the text sample below.

9.42. The order of clitics in the clitic chain was fixed: first the conjunctions *-(y)a* 'and', *-a* 'but', or *-ma* 'but'; then the quotative particle *-wa(r)*, used to indicate that the clause in which it appears is directly quoted speech; the enclitic pronouns (with forms of the 3rd person preceding the other two persons); the reflexive particle *-za*; and finally the still imperfectly understood aspectual or modal particles *-kan* or *-šan* (also *-(a)šta* and *-(a)pa* in the older language).

9.43. The syntax of relative clauses is of some interest. If the relative pronoun is the first word in its clause (or is preceded only by a sentence connective with or without attached clitics), then it is indefinite: *kue* GAL^{HA} *akkuškizzi ta apie=pat ekuzi* 'whichever cups he usually drinks from, he shall drink from those' (KBo 19.74

iv 33'–34'); *nu=mu=kan kuš idaluš memiaš ZI-nu anda n=an=mu* DINGIR^{MES} EGIR-pa SIG₇-ahhanzi šarlanzi 'whichever bad word (is) in my soul, the gods will alleviate it (and) remove it' (KUB 6.45+ iii 46–47). However, if it is preceded by any elements in its clause, it is definite: GU₄-ya=wa=mu *kuin tet nu=war=an=mu uppi* 'And the cow that you promised me, send it to me' (Mašat 75/14 obv. 14–16). As is typically the case in older IE languages, the relative clause precedes the main clause.

Old Hittite text sample

9.44. Below is an excerpt from the Old Hittite Ritual for the Royal Couple, KBo 17.1 (CTH 324), column III, lines 1–13, following the edition of Otten and Souček 1969. This passage is taken from the middle of the ritual; previously prepared items – an eagle and clay figurines representing an army – are here manipulated in the performance of sympathetic magic (somewhat similar to voodoo).

The passage is given below in narrow transcription; broad transcription is used in the notes (where equals signs mark off the clitics). Brackets enclose missing parts of the tablet that have been restored by conjecture; material inside the bracketed sections that is further enclosed in parentheses has been supplied from identical passages in duplicate copies (in this case, the fragments KBo 19.3 and 19.6). The horizontal lines are paragraph dividers in the original.

- [ma-a-a]h-ha-an-da ⁴UTU-uš ⁴ISKUR-aš ne-e-pí-iš te-e[(-kán-na)]
 2 uk-tu-u-ri LUGAL-uš SAL.LUGAL-aš-ša DUMU^{MES}-ša uk-tu-u-ri-eš
 a-š[(a-a)]n-d[u]
-
- ta nam-ma ^{MUŠEN}ha-a-ra-na-an ne-e-pí-ša tar-na-aḫ-ḫi
 4 a-ap-pa-an-an-da-ma-aš-še ke-e me-e-ma-aḫ-ḫi na-at-ta-an ú-uk
 t[(ar-na-a)]h-hu-un LUGAL-ša-an SAL.LUGAL-ša tar-na-aš nu i-it ⁴UTU-i
 6 ⁴ISKUR-ya me-e-m[(i-i)]š-ki ⁴UTU-uš ⁴ISKUR-aš ma-a-an uk-tu-u-ri-eš
 LUGAL-uš SAL.LUGAL-aš-ša QA-TAM-MA uk-tu-u-ri-eš a-ša-an-tu
-
- 8 ú-i-il-na-aš ERÍN^{MES}-an te-eš-šu-um-mi-uš-ša ta-ak-na-a
 ḫa-ri-e-mi tu-uš tar-ma-e-mi ta ki-iš-ša-an te-e-mi
 10 ⁴UTU-uš ⁴ISKUR-aš ka-a-š[(a LU)]GAL-i SAL.LUGAL-ri DUMU^{MES}-ma-aš-ša
^{URU}Ḫa-at-tu-ši
 e-er-ma-aš-me-et e-eš-ḫ[(ar-š)]a-me-et i-da-a-lu-uš-me-et
 12 ḫa-tu-ka-aš-me-et ḫa-ri-[(e-nu-u)]n ta-at a-ap-pa ša-ra-a
 le-e ú-e-ez-zi

(The priest speaks:) "Just as the Sungod (and) the Stormgod, heaven and earth (2) (are) eternal, so may the king and queen and (their) children be eternal!"

(3) Then I release the eagle to the sky, (4) and after it I speak thus: "I did not release it – the king and queen released it. Go! Say to the Sungod and the Stormgod, 'Just as the Sungod and the Stormgod (are) eternal, (7) let the king and queen likewise be eternal!'"

(8) I bury the clay soldiers and the pots into the earth, and I nail them fast. Then I speak as follows: (10) "Sungod, Stormgod, I have just buried the disease, bloodshed, evil, (and) fright of the king, queen, and their sons in the city of Ḫattuša. Let it not come up again!"

9.44a. Notes (through beginning of line 9 only). 1–2. *nēpiš*: ‘sky, heaven’, a neuter *s*-stem from PIE **nebh-es-* ‘cloud’; cp. Ved. *nābhas-* and Gk. *nēphos*. *tēkann=a*: ‘and earth’; *tēkan* is from PIE **dheghōm*, and the clitic conjunction *-a* means ‘and’ and doubles the preceding consonant (as also in the phrase ‘king and queen’ in the next line). After a vowel and sometimes Sumerograms (e.g. line 6) it is written *-ya*. *uktūri*: ‘eternal’, Old Hitt. neuter collective pl. LUGAL-uš SAL.LUGAL-ašš=a: ‘king and queen’; the underlying Hittite is *ḥaššuš* **ḥaššuššarašš=a*. SAL (or MUNUS) is the Sumerogram for ‘woman’. The ending *-ššarašš* comes from PIE **soro-* ‘woman’, which had marginal use as a feminizing suffix (§6.71). *ašandu*: 3rd pl. imperative, ‘let them be’, spelled *ašantu* in line 7.

3–4. *ta*: Old Hitt. sentence connective. ^{MUSEN}*ḫāranan*: ‘eagle’, accus. sing. *n*-stem; cognate with Gk. *ōrn-ith-* ‘bird’ and the second syllable of German *Adl-er* ‘eagle’ (< Middle High German *adel-ar* ‘noble Aar [eagle]’). ^{MUSEN} is the Sumerian determiner for bird-names. *nēpiša*: ‘to the sky’, directive case. *tarnahhi*: ‘I release’, 1st sing. present of the *hi*-conjugation, like *mēmahhi* ‘I say’ in the next line. *āppananda=ma=šše*: ‘And/but (*-ma*) after (*āppananda*) it (*-šše*)’. *kē*: neut. pl., ‘these things’; PIE pronominal stem **kē-* (§7.9). *natta=an*: ‘not’ plus the clitic animate object pronoun referring to the eagle, ‘it’. *āk*: emphatic 1st sing. pronoun, ‘I’ (§9.11).

5–6. LUGAL-š=an SAL.LUGAL-š=a: underlying Hittite is *ḥaššuš=an* **ḥassuššarašš=a* ‘the king and (*-a*) the queen it (*-an*, i.e. the eagle) . . .’ *tarnaš*: ‘released’, 3rd sing. preterite; singular because the subject ‘king and queen’ are treated grammatically as a unit. *it*: ‘go’, suppletive imperative to the verb *pāi-* ‘go’ from PIE **h₁i-dhi* (Ved. *ihī*, Gk. *itiri*). *mān*: ‘just as’, originally a reduced form of *maḥḥanda* (line 1). **UTU-i* **ISKUR=ya*: ‘to the Stormgod and the Sungod’. The determiner *i*, short for Sumerian DINGIR ‘god’, precedes divine names. The phonetic complement *-i* is the Hittite dative singular ending; *-ya* means ‘and’. **UTU-uš* in line 6 is the nominative sing. *mēmiški*: ‘say, speak’, sing. imperative of the iterative of *mēmai-* ‘speak’, here perhaps with an inceptive meaning: ‘begin to speak’.

7–9. *QATAMMA*: ‘likewise’, Akkadian adverb. *uīnaš*: ‘of clay’, genit. sing. The sentence does not begin with a connective because it starts a new section (§9.41). *ERIN^{MES}=an*: ‘troops’; plural in Sumerian (*MES*), but singular in Hittite, hence the accus. sing. ending *-an*. *teššummiušš=a*: ‘and the pots’, accus. pl. *taknā*: ‘into the earth’, directive of *tēkan* (stem *takn-*); the spelling *tāk-na-a* instead of *tāk-na* probably indicates stress on the ending, consistent with the presumed amphikinetic paradigm of the noun; see §6.27. *ḫariemi*: ‘I bury’.

Luvian

9.45. Luvian (or Luwian) is attested in two closely related dialects, Cuneiform and Hieroglyphic. The language was spoken over a wide territory, especially to the south of the Hittites but also to the west. Evidence has been mounting that the fabled city of Troy, on the northwest coast of Turkey, was Luvian-speaking (see further below). As the Hittites spread south and west over Anatolia, they became strongly influenced by Luvian culture and borrowed many words, especially in the domain of religion and cult. Luvian was employed in rituals adopted by the Hittites, and most preserved Cuneiform Luvian is embedded in Hittite ritual texts from the Middle and Neo-Hittite periods.

Hieroglyphic Luvian is the only Anatolian dialect attested in both the second and first millennia BC. It is written in a native hieroglyphic script that was developed by the fourteenth century BC, although during the Hittite Empire was mostly limited to writing proper names on seals. Hieroglyphic Luvian inscriptions begin to appear on stone faces and rocks starting around the thirteenth century BC, but are infrequent before the destruction of Hattusas shortly after 1200 BC, after which Hittite ceased to be the official language of culture. The Luvian of the hieroglyphs then filled this

role – a shift that in some respects was not sudden: already in the later Neo-Hittite texts there is a steady increase of Luvianisms. The majority of Hieroglyphic Luvian inscriptions date from the ninth to the seventh centuries BC, although some of the dates are being revised and may be earlier. The inscriptions include historical accounts and economic records.

9.46. Both forms of Luvian present certain difficulties. Cuneiform Luvian script is the same as Hittite script, but the meanings of many words are still opaque. The hieroglyphic script, like cuneiform, consists of characters used both logographically and syllabically; unfortunately, it is phonetically much less highly developed than cuneiform. The syllabic signs only stand for single vowels or CV sequences, and no effort is made to distinguish long vowels or many other basic phonetic facts. Since the hieroglyphic script gives so little information about the pronunciation of the language, almost all our phonetic knowledge has been gleaned from Cuneiform Luvian (the two varieties of Luvian are phonologically nearly identical). Major breakthroughs in the past few decades have improved the readings of a number of hieroglyphic signs, but some remain undeciphered.

It is widely thought that Lycian is a descendant of a variety of Luvian; see below.

The Trojan connection

9.47. Evidence has gradually been growing that the Trojans spoke an Anatolian language, perhaps Luvian. Troy is now generally identified with an Anatolian city called Wilusa (*Wiluša*), whose name is strongly reminiscent of *Wilion*, the early form of *Ilion*, a Greek word for Troy. One of the kings of the Hittite Empire, Muwatallis, is remembered for the treaty he sealed with one Alaksandus of Wilusa; the name is probably to be identified with Aleksandros, the Greek name for the Trojan prince Paris. In one passage of Luvian poetry, Wilusa is called 'steep', an epithet also used of the city by Homer. In 1995 a seal inscribed in Luvian was unearthed on the site of Homer's Troy, the first written object to be found there; though tantalizing, this does not yet prove that the Trojans spoke Luvian.

Developments from Common Anatolian to Luvian

Grammatically, Luvian is very similar to Hittite, but there are some differences in the phonology and morphology.

PHONOLOGY

9.48. Luvian (and perhaps Lycian) appears to be the only IE language that preserves three separate outcomes of the three PIE velar series, at least when voiceless. As in Hittite, the plain velar **k* and the labiovelar **kʷ* remain intact: *kīšammi* 'combed' < **kes-* (cp. OCS *češō* 'I comb'); and *kuiš* 'who' < **kʷis* (cp. Lat. *quis* 'who'). Unlike Hittite, **k* remained distinct from **k* and became *z* (pronounced *ts*, as in *ziyari* 'lies' < **kei-o-*, cp. Ved. *śāye* 'lies' < **kei-o-*). The voiced velars did not develop in parallel with the voiceless velars: **gʰ* before a front vowel disappeared, as in *iššara* 'hand' < **gʰes-ro-* (cp. Hitt. *kišsar-*, Gk. *kheir* < **gʰes-r-*); and **gʷ* became *w*, as in the word for 'woman', *wāna-* < **gʷen-*.

9.49. A characteristic feature of Luvian phonology is a change known as Čop's Law, after the Slovenian Indo-Europeanist Bojan Čop. Under this rule, an original

stressed *é before a voiced stop, resonant, or s became a, and the following consonant is written double. (The consonant may have undergone fortition or “strengthening” of some sort, such as devoicing or doubling.) Thus PIE *médhu ‘sweet drink’ became *maddu* ‘wine’; *mélit- ‘honey’ became *mallit-*; and *yé-s- ‘good’ shows up in *wašsar-* ‘favor’. A limited version of this change may have already happened in Common Anatolian, but this is uncertain.

9.50. Lenition. Luvian and Lycian provide almost all the evidence for a set of lenition (“weakening”) rules in Common Anatolian, which voiced an intervocalic stop after a long stressed vowel or when between two unstressed vowels. Orthographically, such stops were written single, rather than double as was usually the case with the reflexes of voiceless stops. The lenition also affected laryngeals, which are written as single in such positions rather than double; this could also indicate voicing. It has been proposed that this lenition is of Common Anatolian date, but this is controversial; the best evidence for the lenition rules comes from Luvian and its probable descendant Lycian. Thus Luvian *malliati* ‘with honey’ (instr. sing.) has single *t*’s following original unstressed vowels (earlier **mélitoti*); compare also Lycian *tadi* ‘puts’ (< Common Anatolian **dēti* < PIE **dbéh₂ti*).

9.51. Differences between Cuneiform and Hieroglyphic Luvian. As briefly noted earlier, Hieroglyphic Luvian differs from Cuneiform Luvian only in minor detail. The words for ‘heaven’ reflect different ablaut grades: Cuneiform *tappaš-* from **nebes-* (with Čop’s Law) and Hieroglyphic *tipas-* from **nēbbes-*; the original paradigm was an acrostatic *s*-stem, nomin. sg. **nēbb-os*, weak stem **nēbb-es-*. But both languages share the innovation of changing the initial **n-* in this word (cp. Hitt. *nēpiš-*, OCS *nebes-*) to a dental stop (interestingly paralleled by Lith. *debesis* ‘cloud’). Aside from such differences in vocabulary, Hieroglyphic Luvian sometimes rhotacized **d* to *r* between vowels, as in *pa+ra/i-* (phonetically *para-*) ‘foot’ < **podō-*.

MORPHOLOGY

9.52. The noun has some peculiarities from the Hittite point of view. There was no genitive case in Cuneiform Luvian; to show possession, the language instead turned the noun into a relational adjective having the suffix *-ašša-*. Thus the noun *tiyamm(i)-* ‘earth’ formed the relational adjective *tiyammašša-* ‘pertaining to the earth, of the earth’. Such relational adjectives are not exact equivalents of genitives, since they cannot distinguish between the singular ‘of the X’ and the plural ‘of the X-es’; Luvian solved this problem by adding the relational adjective suffix to an old animate accusative plural ending **-anza* if the base noun was plural. Thus the word for ‘ritual’ was *malḥašša*, which always declined in the plural; its relational adjective was *malḥašš-anz-ašša-* ‘pertaining to the ritual’.

9.53. Neuter nouns in the nomin.-accus. singular regularly appear with a particle *-ša* attached, quite probably from an old neuter demonstrative **sod*. An example is *wār-ša* ‘water’ (< **uoh₂r*, cp. Ved. *vār*). As elsewhere in Anatolian, neuters are declined in an ergative case when functioning as the subject of a transitive verb.

9.54. Another peculiarity of Luvian nominal declension not found in Hittite is that certain animate nouns and adjectives formed their nominatives and accusatives with a stem *-i-*, even if they were not normally *i*-stems. Thus the relational adjective *tiyammašša-* ‘of the earth’ above has a nominative sing. *tiyammaššiš*. This is also the norm in Lycian and Lydian. The origins of this phenomenon (termed “*i*-mutation”) are debated.

9.55. In verbal morphology, Luvian differs from Hittite especially in the 1st person present active endings. The singular normally ends in *-wi*, as in *ḥapiui* 'I bind' and *arkamanallau* 'I bear tribute'; and the plural is *-unni*, related to Hitt. *-wani* (e.g. *mammalḥunni* 'we crush'). Luvian has an unexplained 3rd plural preterite ending *-aunta* (e.g. *nakkušāiunta* 'they furnished a scapegoat', Hieroglyphic Luvian *wali-la-u-ta* 'they died', phonetically *walaunta*). Luvian kept **t* intact before *i*, unlike Hittite, so the 3rd sing. and pl. of the *mi*-conjugation are still *-ti* and *-nti*: *iti* 'goes' (cp. Lat. *it*), *ḥišḥiyanti* 'they bind' (vs. Hitt. *išḥiyanzi*). Because of the limited remains of the language, not all forms of the verbal paradigm are known.

9.56. Not much is known about the non-finite formations in Luvian. It has a participle in *-mma-* (or *-mmi-*) that functions like *-ant-* in Hittite or **-tō-* elsewhere in IE: *kīšammi-* 'combed', *dūpaimmi-* 'beaten'. The source may be PIE **-mo-* that yields passive participles in Balto-Slavic (§5.56), but this is disputed. (This feature is shared by Lycian.) An infinitive in *-una* is also found (e.g. *aduna* 'to eat'), probably an old directive case of a verbal noun in **-uar*, stem **-un-*.

Luvian text sample

9.57. From the Dūpaduparša Ritual, KUB 9.6 + 35.39 i 23f. The ritual is written mostly in Hittite, with spells and incantations in Luvian spoken by the "old woman" (^{SAL}ŠU.GI) who officiated over the ritual. Restorations have been made on the basis of parallel passages elsewhere on the tablet. The interpretation of several forms is tentative. An 'x' stands for an unreadable sign.

- ku-iš ḥi-i-ru-ta-ni[-ya-at-t]a ti-wa-ta-ni-ya-at-ta
 24 na-a-nu-um-pa-ta ma-a[d-du-ú]in-zi ma-al-li-ti-in-zi
 da-a-i-ni-in-zi x[]x-al-la-an-zi a-ar-ši-ya-an-du
 26 [t]a-a-i-in-ti-ya[-ta ma-al-li a-i-]ya-ru ta-pa-a-ru-wa
 [ḥi-ru]ú-ta [ta-ta-ar-ri-ya-am-na u-w]a-la-an-te-ya
 28 ḥu-u-i-it-wa[-li-e-ya]ḥi-e-ya na-a-ni-e-ya
 na-a-na-aš-ri[-e-ya]
 30 lu-u-la-ḥi-e-ya ḥ[a-pí-ri-e-ya ku-wa-ar-š]a-aš-sa-an
 tu-ú-li-ya-aš-ša-a[n]

Whoever has cursed (or) sworn, (24) then now let the . . . of wine, of honey, (25) of oil flow! (26) Let them become oil (and) honey – the . . . s, (27) the oaths (and) curses of the dead (28) (and) of the living, of the . . . , of a brother, (29) of a sister, . . . (30) of the mountain-dwellers, of the people of the plain, of the army, (31) of the assembly . . .

9.57a. Notes. 23. *kuiš*: 'who, whoever', PIE **k^wis* (Hitt. *kuiš*). *ḥirutaniyatta*: 'has cursed', preterite 3rd sing., to be analyzed as *ḥirut-aniya-tta*, a denominative verb from the noun *ḥirut* 'curse'. The ending *-tta* is the 3rd sing. preterite ending, from PIE **-to*, and *-aniy(a)-* is a denominative suffix that ultimately comes from PIE **-je/o-* (§5.30). The noun *ḥirut* 'curse' (nomin. pl. *ḥirūta* below in line 27) comes from **h₂ērū-t-* (with long *ē* not colored by the laryngeal) and is probably cognate with Gk. (Arcadian) (*kat-jaru(-os)* 'accursed', *tiwataniyatta*: 'has sworn, cursed', a 3rd sing. preterite denominative like the preceding, from *tiwat-* 'Sungod' (**diu-ot-*, like Lat. *diu-it-* 'wealthy'). It has been suggested that it originally meant something like 'call upon god' in a bad sense, like the Oscan verb *deuma-* 'curse' from the word for 'god'.

24–25. *nānum=pa=ta*: a chain consisting of *nānun* ‘now’ plus a conjunction *-pa* plus a particle *-ta* that seems to mean ‘then’ or the like. *maddūinzi*: ‘of wine’, nomin. pl. of *madduwi(ya)*-, adjective formed from *maddu* ‘wine’, PIE **medhu-*, with geminate *-dd-* and *-a-* by Čop’s Law (§9.49 above). *mallitinzi*: ‘of honey’, nomin. pl. of *malliti(ya)*-, from *mallit-* ‘honey’, PIE **melit-*, with geminate *-ll-* and *-a-* by Čop’s Law. *dāininzi*: ‘of oil’, nomin. pl. of *daini(ya)*-, from *dain-* ‘oil’. *āršiyandu*: ‘let (them) flow’ or ‘let (them) make them flow’, 3rd pl. imperative; cp. Hittite *aršiya-* ‘cause to flow’.

26–27. *tāin=tiy=ata*: ‘oil’ (nomin. sing.) plus the reflexive particle *-ti* plus the clitic pronoun *-ata* ‘they, them’, with a *y*-glide written between them. The clitic pronoun “doubles” the subject and is required when the subject (*tapāruwa hirūta* and so forth) is shifted to the end of the sentence (technically, when it is right-dislocated). *aiyaru*: ‘let become’, 3rd sing. imperative; singular verb agreeing with a neuter plural subject, an archaic IE syntactic feature (§§6.68, 8.16). *tapāruwa*: of unknown meaning, but apparently designating something bad; the first in a string of neuter plurals. *tatarriyanna*: ‘curses’ (the restoration is tentative); lit. ‘(bad) things that get said’, from *tatarriya-* ‘say’ (cp. Hitt. *tar-anzi* ‘they say’). *uwalanteya*: ‘of the dead’, neut. pl. of a relational adjective in **-ejo-*, as are all the following words. The base *uwalant-* ‘dead’ is perhaps the clearest Luvian example of an old participle in *-nt-*, §9.4.

28–31. *hūitwalieya*: ‘of the living’, from *hūitwal-* ‘alive’, related to Hitt. *hūišu-*, *hūišwant-* ‘living, alive’, with unclear morphological details. *nānieya*: ‘of a brother’, adjective from an unattested **nami-* that would be the Luvian cognate of Hitt. *negnaš* ‘brother’ (interestingly not the PIE word). *nānašrieja*: ‘of a sister’, from the ‘brother’ word plus the feminizing suffix *-šri-*; see note to line 2 of the Hittite text above. *lūlaḫieja*: ‘of those living in the mountains’. *ḫapirieja*: ‘of those living on the plain’. *kuwaršašsan*: ‘of the army’; apparently the *-an* in this and the following word is a mistake for *-a*, since they should both still be neuter plural. *tūliyašsan*: ‘of the assembly’. The rest of the line is unclear.

Palaic

9.58. Palaic is known from only eleven ritual and mythological texts written down during Old Hittite times. The language seems to have become extinct rather early, and has in fact been called the first Indo-European language to die out. The name comes from Palā, the Hittite name of a region probably in north-central Anatolia across the Halys River in the area known in classical times as Paphlagonia (a name plausibly interpreted as deriving from a reduplicated version of Palā, **pa-pla-* or the like).

9.59. Palaic is on the whole quite similar to Hittite. A striking phonological feature is the continuation of the sequence **-eh₂-je-* in derived verbs as *-āga-*. As in Hittite and Luvian, short accented vowels in open syllables are written long, as in the genitive singular of the word for ‘eagle’, *ḫāranāš* < **h₂óron-*; contrast the nomin. sing. *ḫarāš=kuwar=zi* in the text below, where the accent seems to have shifted to before the following enclitic *-kuwar*.

Palaic text sample

9.60. KUB 32.18 i 5’ff. (The symbol ‘ means the top part of the tablet is broken off and lines have been numbered from the beginning of the surviving part of the tablet.) The text is too poorly understood to provide a complete translation. Line 7’ is essentially identical to a passage from the Hittite Telipinu myth (KUB 19.10) which states, *eter n=e UL išpier ekuier=ma n=e=za UL ḫaššikir* ‘They ate, (but)

they were not satisfied; they drank, (but) they were not satisfied.” This seems to be a mythological theme of Common Anatolian date. An ‘x’ stands for an unreadable sign.

- 6' |x-na-ku-pa-an-ta šu-wa-a-ru' ša-a-ú-i|
 |an-za ma-a-ar-za ma-a-aḥ-la-an-za a-an-ti-en-ta ma-a[-ar-ḥa-aš
 [a-]ta-a-an-ti ni-ip-pa-ši mu-ša-a-an-ti a-ḥu-wa-an-ti ni-ip-pa-aš ḥa-ša-an-ti
-
- 8' [ti-]ya-az-ku-wa-ar ú-e-er-ti ka-a-at-ku-wa-a-at ku-it a-ta-a-an-ti
 [ni-]lip-pa-ši mu-ša-a-an-ti a-ḥu-wa-a-an-ti ni-ip-pa-aš ḥa-ša-a-an-ti
- 10' x ḥa-ra-a-aš-ku-wa-ar-zi pa-na-a-ga-an-zi ši-i-it-tu-wa-ra-an
 [ši-]jit-ta-an ḥa-pi-it-ta-la-an-ku-wa-ra-an ši-it-ta-an

9.60a. Notes. 5'–7'. *šuwāru*: 'full', neuter nomin.-accus. sing., cp. German *schwer* 'heavy', Lith. *svarūs* 'heavy'. The raised exclamation point is a convention indicating that the scribe wrote a different sign by mistake. *šau*i[: probably the word for 'horn', whose collective plural is *šawidār*; cp. Hitt. *šawatar* 'horn'. Note that if this interpretation is correct, we have a reference here to a 'full horn' or cornucopia. *mārza*: something edible. *āntienta*: 'they go in', if it is to be segmented *ānt-yenta* (cp. Hitt. *anda* 'in, into' and Hitt. *ianzi, ienzi* 'they go' < root **h₂ei-* 'go', cp. Lat. *eunt* 'they go'). *mārhaš*: 'the gods'. *atānti*: 'they eat', Hitt. *atanzi*, PIE **h₂edonti*. Note the scriptio plena in the ending *-ānti* in this and the next verb (and more consistently in lines 8' ff.), which probably indicates stress. *nippaši*: to be segmented *ni=ppa=ši*, of which *ni-* is the negative 'not', *-ppa* is a conjunction meaning 'but', and *-ši* is probably a reflexive particle (note that the Hittite word for 'eat one's fill', *ḥaššikke-*, is also used with a reflexive particle). *mušānti*: 'they eat their fill, become satisfied'; cp. Greek *amutí* 'at one draught'. *aḥuwanti*: 'they drink', Hitt. *akuwanzi*; Palaic lenited (weakened) the intervocalic labiovelar to *-ḥu-* (PIE **eg^wh-* 'drink', also in Lat. *ēb-rus* 'drunk'). *nippaš*: *ni=pp(a)=aš* 'but they not'; *-aš* is probably a 3rd pl. subject pronoun. *ḥašanti*: 'they drink their fill, become satisfied', cp. Hitt. *ḥaššikke-* 'eat one's fill'.

8'–11'. *tiyaz*: 'the Sungod', PIE **diyōts*, cp. Luv. *Tuwaz*. The particle *-kuwar* may be emphatic: 'the Sungod himself'. *uērti*: 'announces, says'; cp. Hitt. *weriya-* 'say'. *kātkuwāt*: *kāt* means 'this', nomin.-accus. sing. neuter (or maybe plural); the rest is unclear. *panāganzi*: 'appearing' or the like; animate nomin. sing. present participle, with the *-i* probably an empty vowel, from **-ant-s* (in Hittite, this would have been written *-anza* with an empty *-a*). The stem goes back to something like **panāh₂ie-*, with the Palaic outcome of **-ah₂ie-* discussed above. *šitruwaran šittan*: unclear in meaning, but probably a middle verb form *šittuwar* (plus clitic pronoun *-an*) followed by a 2nd pl. imperative (active) *šittan* from the same verb. The alternation in voice is strongly reminiscent of Cuneiform Luvian *azzaštan* . . . *niš aztiwari* 'Eat! (2nd pl. imperative active) . . . You do not eat' (2nd pl. middle) in KUB 9.31 ii 26–28. *ḥapitalan*: 'member, limb, joint'; cp. Hitt. *ḥapp-eššar* 'limb, joint' with different suffix.

Lycian

9.61. Lycian was spoken in Lycia, a country on the southwest Anatolian coast in a rugged area between the Gulf of Fethiye and the Gulf of Antalya. It is attested almost solely on coin legends and short tomb inscriptions from the fifth and fourth centuries BC in a Greek-derived alphabet. Only two extensive texts have been found, the Xanthos Stele, containing a historical account, and the Letoon Trilingual, concerning the establishment of a cult of Leto and written in Lycian, Greek, and Aramaic. The Trilingual, not surprisingly, has been crucial for understanding the

language. The Lycian on two sides of the Xanthos Stele is different from the usual sort and is called *Milyan* or Lycian B (with Lycian A being the ordinary form of Lycian). As stated above, it is usually thought that Lycian continues a variety of Luvian, though not the same one that has come down to us.

9.62. The letters used in the traditional transcription of the Lycian alphabet do not all match their probable pronunciations very well. The letter transcribed as χ was not a fricative but a stop; what are written as voiced stops were pronounced as fricatives (so, e.g., d stands for δ); and z represents a voiceless affricate ts as in Hittite. Additionally, after nasals or nasalized vowels, it appears that stops were voiced even if they were written voiceless. For example, the possessive suffix *-wāt(i)-* (PIE **-uent-*, §6.40) was pronounced *-wād-* (with nasalized a). Etymological d 's show up spelled as t 's in this position too, as in *ñte* 'in' < **endo* (cp. Hitt. *anda*, Gk. *endo-*).

9.63. As might be expected of a language descended from or otherwise closely allied with Luvian, the three original voiceless velar series have distinct reflexes. The old palatal velar **k̥* became s , as in the verb *si-* 'lie' from **k̥ei-* (cp. Luvian *zi-* seen above, §9.48). The ordinary velar **k* remains k in all positions, but **kʷ* becomes t in Lycian A before front vowels, as in *ti-* 'who, which' < **kʷi-* 'who, which' (compare §12.15 for the same development in Greek). Our information on the voiced velars is meager, but from what we have it appears that they pattern with Luvian. For example, **gʷ* also became w , as in *waiva-* 'cow' < **gʷōu-*.

9.64. Consonantal reflexes of the second laryngeal are well attested; **h₂* became a stop variously spelled k , q , and χ , with the precise conditioning factors still not worked out. Examples include *χñt-* 'front' (cp. Hitt. *hant-*) and *Trqqas* (stem *Trqqñt-*), the name of the Stormgod (cp. Luv. *Tarhunt-*) < **trh₂uent-*. The latter example suggests that q is a labialized outcome of the laryngeal before the glide **u*.

9.65. The fate of vocalized laryngeals in Anatolian is the subject of controversy. Figuring in the debate is a famous word in Lycian A, *kbatra-* 'daughter' (PIE **dhugh₂ter-*), which most likely preserves the **h₂* as the first *-a-*. An early **dugatr(a)-* lost its internal g , becoming **duatra-*; in Lycian A, the cluster **d̥u-* became $kb-$ (compare the word for 'two', *kbi-* < **d̥ui-*), so **duatra-* then became *kbatra-*.

9.66. Unaccented vowels in various positions underwent syncope (loss), although the exact conditions still have to be worked out: *ahñtai* 'property' < **ahāntai* < **asant-* 'existing'; *pddē* 'place' < **pedom*; *ōōē* 'votive object' < **dVhē* or **iVhē*. As these last two examples show, this syncope produced various unusual word-initial consonant clusters, giving Lycian part of its characteristic look. The final consonant in consonant clusters is often written double, as in both *Trqqas* and *pddē*; the reasons for this are not fully clear.

9.67. Lycian provides crucial evidence for the existence of the vowel **o* in Common Anatolian, for it is the only language in which this vowel did not merge with *a* either in pronunciation or in spelling. (As the cuneiform syllabary had no signs for *o*, we cannot be absolutely sure that *a* was not sometimes used to write the vowel *o*.) Its outcome in Lycian was e , as in the preterite 3rd sing. verbal ending *-te* from **-to*, the neuter nomin.-accus. ending *-ē* from **-om*, and words like *ñte* 'in, into' < **endo*; contrast for all of these Hittite and Luvian *-ta*, *-an*, and *anda*.

9.68. Lycian syntax differs from the syntax of its Anatolian sisters in not having verb-final clause structure. Verbs appear near the front of the clause, preceded by topicalized or focused elements and certain conjunctions. Apparently, the option of fronting verbs for emphasis (§8.19) was generalized in this language.

Lycian text sample

9.69. The beginning of the Letoon Trilingual. The translation below loosely follows an unpublished one by H. Craig Melchert (used here by permission). Each clause is put on a separate line below for clarity.

- ēke Trñmisiñ xssaθrapazate Pigesere Katamlah tideimi
 sē=ñine ñite-pddē-hadē Trñmile pddēnehñmis Ijeru se-Natrbbijēmi
 sej-Arñna asaylazu Erttimeli
 me-hñti-tubedē aruz sej-epewētlñmēi Arññāi
 5 ññmaitē kumeziñē 00ē xñtawati xbidēñni sej-Arkazuma xñtawati
 sē=ñin=aitē kumazu mahāna ebette Esemiju Qñturahahñ tideimi
 se=de Esemijaje xuwati-ti

When Pigesere (Pixodaros) son of Katamlē (Hekatomnos) ruled as satrap of Lycia (2) and he installed for the Lycians Hieron and Natrbbijēmi (Apollodotos) as commissioners (3) and for Xanthos Artemelis as governor, (4) the free citizenry (?) and the dependents of Xanthos came to an agreement (5) and they made a sacred cult installation for the (gods) King of Kaunos and King Arkazuma. (6) And they made Simias son of Kondorasis priest for these gods, (7) and the ones who succeed Simias.

9.69a. Notes (through line 5 only). 1. ēke: 'when'. Trñmisiñ: 'Lycia', accus. sing. xssaθrapazate: 'ruled as satrap', 3rd sing. preterite. A satrap was a local governor in ancient Persia; it is ultimately an Iranian word (Old Pers. *xšaθrapāvā* 'protector of a province'). tideimi: 'son', perhaps from a mediopassive participle of a verb **tid(e)i-* 'suckle'.

2. sē=ñine: lit. 'and (se) him (-ē) them (-ñine)', clause-initial clitic chain. The clitic object pronoun -ē 'him' (< *-om, cognate with Hitt. -an) doubles the direct object *Ijeru*, and the clitic pronoun -ñine 'them' doubles the indirect object *Trñmile*. Such clitic doubling is obligatory in Lycian syntax. ñite-pddē-hadē: 'installed, commissioned'; ñite- is a preverb cognate with Hitt. *anda* 'in, into'; pddē is accus. sing. of the noun meaning 'place' and seems to form part of a compound verb, the other part being *hadē* 'let go of, left'; literally then 'left in place'. Trñmile: 'for the Lycians', dat. pl.; cp. *Trñmisiñ* in the preceding line. pddēnehñmis: 'commissioners', accus. pl. The second half of it, -hñmi-, is the mediopassive participle of *ha-* 'let, leave' seen above. *Ijeru se-Natrbbijēmi*: 'Hieron and Natrbbijēmi'; the first is a Greek name, the second is translated in the Greek version of the inscriptions as Apollodotos ('given by Apollo'), meaning that *Natr-* is an Anatolian god corresponding to Apollo, and -bbijēmi is *pijēmi* 'given' (with *p-* voiced to *b-* following the -r of *Natr-*), the participle of *pije* 'give' (cp. Hitt. *pāi* 'gives').

3-5. sej-Arñna: 'and for Xanthos', dat. pl. (with singular meaning); the -j of *sej* is just a hiatus-filling glide between vowels. asaylazu: 'governor'. Erttimeli: 'Artemelis', Greek name. me-hñti-tubedē: 'they made an agreement'; me- is a conjunction that introduces a main clause following a dependent clause; hñti- may consist of *hñ-* 'together' plus the reflexive particle -ti (Hitt. -za, Luv. -ti). aruz: uncertain, here rendered as 'citizenry'. sej-epewētlñmēi: 'and the *perioikoi*', the neighboring dependent population; ultimately a participle of an unknown verb. Arññāi: 'of Xanthos', genit. pl. ññmaitē: a verb form of uncertain meaning. kumeziñē: 'sacred', derived from *kumaza-* 'priest' with the same "occupational" suffix -aza- as in the word for 'governor' in line 3. 00ē: 'cult offering, installation'; cp. Lydian *tasēv* 'stele'. xñtawati: 'king', dat. sing.; probably literally 'the one who is in front/foremost', cp. Hitt. and Luv. *hant-* 'front'. xbidēñni: 'of Kaunos', a deity. sej-Arkazuma: 'and Arkazuma', dat. sing.; a name or title of a deity, with *kk* representing some kind of rare velar rendered in Greek by a *g*.

Lydian

9.70. Lydia, located on the west-central Anatolian coast, was inhabited in the first millennium BC by an Anatolian people who left behind a little over 100 texts in an alphabet similar to eastern varieties of the Greek alphabet, as well as in the Old Phrygian alphabet. The texts are mostly from the fifth and fourth centuries BC in Sardis, the capital of Lydia. A few others are dated to as early as the eighth century BC. The inscriptions, carved on stone, include dedicatory formulas, decrees, and epitaphs. The language is difficult to figure out due to the small size of the corpus – the smallest in Anatolian after Palaic. A number of the inscriptions are in verse, which has allowed an analysis of the language's stress placement.

9.71. As in Lycian, stops in Lydian were voiced after nasals. There were two sibilants, transcribed *ś* and *s*; confusingly, it now appears that the first one stands for the ordinary dental or alveolar sibilant *s*, while the second stands for a palatal sibilant of some kind. The symbol *λ* is used to represent what was probably a palatalized *l*. The pronunciation of the symbol transcribed as a lower-case Greek letter nu, *ν*, is uncertain; its only clear source is a word-final nasal.

9.72. Lydian underwent loss of unstressed vowels, as in *alarmś* 'self' < **-mos*, *bidv* 'I gave' < **piyom*, and *ētamv* 'designation' < **ētamēv*. The nasalized vowels *ā* and *ē* occur as in Lycian. An interesting feature of Lydian verbs is that the present 3rd person forms are the same in the singular and plural, and the preterite 3rd person shows an ending *-l*.

9.73. Not much of the morphology of Lydian has been preserved, but that which has survived is consistent with the other Anatolian languages. In nouns, the animate nomin. sing. ended in *-ś* or *-s*, essentially unchanged from PIE, while neuter nouns and adjectives end in *-d*, which has spread from the pronominal neuter nomin.-accus. ending **-d*. As in Luvian, there was no genitive. Possession was expressed by a relational adjective in *-l*, as in *bakillā* 'of Bacchus', or (less commonly) *-si-* (etymologically the same as Luvian *-ašša-*).

Lydian text sample

9.74. The Lydian-Aramaic bilingual epitaph from Sardis, on a marble stele dated to about 400 BC. Parts of the Aramaic version are not well understood, hence some of the uncertainty in the translation. The Lydian version is probably missing a line at the beginning.

[o]raλ isλ bakillā est mruδ eśś=k [wānaś] laqrisa=k qela=k kud=k=it ist eśλ wān[aλ] bλtarwod ak=ad manelid kumlilid silukalid ak=id n[āqis] eśλ mruλ buk eśλ wānaλ buk esvav laqirisav buk=it kud ist eśλ wānaλ bλtarwod ak=t=in nāqis qelλk fēnsλifid fak=mλ artimuś ibśimsis artimu=k kulumsis aaraλ biraλ=k kλidaλ kofuλ=k qiraλ qelλ=k bilλ wcbaqēnt

... during the month of Bacchus. This stele and this [tomb] and *laqrisa* and land (?), (and?) as it ... from (?) this tomb, it (is) of Manes (son) of Kumlis (son) of Silukas. But (?) should anyone (be?) at this stele or at this tomb or at these *laqrisa* or indeed as it belongs in this tomb (??), should anyone violate against the land,

Artemis of Ephesus and Artemis the Koloan will trample (?) on his yard and house, earth and water, property and land.

9.74a. Notes. *orał*: 'month'. *isla*: some sort of participle. *bakilla*: 'of Bacchus'; see §9.73. est: 'this', neuter; the *-t* is a suffixed neuter nomin.-accus. ending ultimately from PIE **-d*. The pronoun *es-* has been plausibly etymologized as cognate with Hitt. *asī* 'the aforementioned'. *mrud*: 'stele', shortened form of *mrūwaad*. *eś=k*: 'and this'; *-k* is the conjunction 'and' < PIE **k^we*. *wānaś*: 'tomb'. *laqrīsa*: part of the tomb or tomb complex. The spelling *laqrīsa* further in the inscription is an error or reflects epenthesis. *qela*: usually interpreted to mean 'land', but uncertain. *kud=k=it*: *kud* is apparently a conjunction, perhaps 'as'; *-it* is 'it'. *ist*: perhaps 'from', from **eś* (cp. Lat. *ex* 'out of, from'). *eśla wānaś*: 'this tomb', object of *ist*. *bātarwod*: a verb of unknown meaning. *ak=ad*: *ak* is apparently a sentence connective like Hittite *nu*; *-ad* is an enclitic neuter pronoun. *manclid*: 'of Manes', relational adjective in the neuter, modifying *-ad*; the following two words are relational adjectives also, functioning as patronymics. *nāqis*: 'whoever, anyone', from earlier **nāvqis*, with the same morphemes (in reverse order) as Latin *quisnam* 'whoever'. *buk*: 'or'. *esvav*: 'these', genit. pl.; since the nasal transliterated as *nu* (v) seems to come from original word-final nasals (§9.71), the original form was **esom*, which, after syncope to **esv*, was recharacterized with the productive ending *-av*. *fēnsλifid*: perhaps 'violates'. *fak*: a variant of *ak* (for the *f-*, cp. Luv. and Palaic *-pa* 'but, and'). *artimūś ibśimsis*: 'Artemis of Ephesus', local goddess. *aarał*: 'yard'. *birał*: 'house', cp. Hitt. *pēr* 'house'. *kāidał*: 'earth'; a suggested cognate is Eng. *clay*. *kofuł*: 'water', perhaps cognate with Armenian *cov* 'sea'. Note the rhyming pair *aarał birał* and the alliterative pair *kāidał kofuł*, poetic features. *qirał*: 'property', perhaps specifically immovable property. *qel=k bīla*: the Aramaic here has 'and whatever is his', which would mean that *qel* is not a form of *qela* 'ground' but rather a form of *qeśi* 'whatever'. *wcbaqēnt*: perhaps 'trample', but uncertain. Lycian does not distinguish between 3rd sing. and 3rd pl. verbs.

Carian, Pisidian, and Sidetic

9.75. As mentioned at the outset of this chapter, three very fragmentarily attested languages of southwest Anatolia also belong to the Anatolian branch: Carian, Pisidian, and Sidetic.

Caria was a region in the extreme southwest of Anatolia, to the north and west of Lycia. A few inscriptions from about the fourth and third centuries BC have been found there, but most of the Carian corpus comes from Egypt, where a number of Carian immigrants settled and left epitaphs and graffiti dating from the seventh to the fourth centuries BC. Recent advances in the decipherment of the Carian alphabet have allowed some preliminary grammatical analysis. The claim that Carian is Anatolian was originally based on the presence of relational adjectives of the Luvian type, and with a related suffix *-ś*. To this can be added some cognates with other Anatolian languages, like *xi* 'who' (cp. Hitt. *kuiš*, and recall also §8.28) and *ted* 'father' (cp. Lyc. *tede-*, Cuneiform Luv. *tātīš*).

Pisidia, located east of Caria and north of Lycia, is the site of a little over 30 tomb inscriptions from perhaps the third or second centuries BC in a Greek-derived alphabet. Finally, Sidetic, the least well attested Anatolian language (if it is Anatolian), is known from a half-dozen inscriptions from the third century BC found in Side, on the Pamphylian coast. The identification of both Pisidian and Sidetic as Anatolian rests mostly on the presence of relational adjectives in *-s*. Note also Sidetic *mašara*, apparently meaning 'gods', which is related to Lycian *mahāna* (which appeared above in §9.69, line 6).

For Further Reading

Most of the important reference works on the Anatolian languages are in German. An exception is the excellent new historical phonology by H. Craig Melchert (1994). Fundamental for Anatolian studies are good textual editions; these are scattered among innumerable journals and other publications, but a number of important ones can be found in the 40-plus volumes of the ongoing series *Studien zu den Boğazköy-Texten* (StBoT for short). This series also contains the complete Cuneiform Luvian corpus (Starke 1984) and the complete Palaic corpus with glossary (Carruba 1970). Two etymological dictionaries of Hittite are currently in progress, one in English (Puhvel 1984–, with copious textual citations) and one in German (Tischler 1977–). A thorough grammar of Hittite written in English, with notes and exercises, is in preparation by Harry A. Hoffner, Jr. and Craig Melchert, and will supplant Friedrich 1974, still the standard reference grammar. Melchert 1993 is a complete dictionary of Cuneiform Luvian, with etymological notes. The first-millennium-BC Hieroglyphic Luvian corpus, except for one text that will be separately published, has appeared in an exquisite new edition (Çambel 1999 and Hawkins 2000). The only grammar of Lycian, with texts, is Neumann 1969; the Lydian corpus and grammar are contained in Gusmani 1964 (plus supplements). A complete Lycian dictionary, with etymologies, is now Melchert 2004.

A comprehensive book on Anatolian historical morphology is still lacking; Kronasser 1956 is about the closest thing to one, but it has little on the non-Hittite languages and contains many out-of-date views. For athematic nouns in Hittite, see now Rieken 1999 (see Bibliography, ch. 6), and for Luvian, see Starke 1990, though not all his views have been followed. Anatolian syntax took a large leap forward with Garrett 1990 (see Bibliography, ch. 8), a pathbreaking doctoral thesis; he has since written additional articles on the subject.

Emmanuel Laroche's *Catalogue des textes hittites* (see §9.20) is being periodically updated electronically at <http://www.asor.org/HITTITE/CTHHP.html>.

For Review

Know the meaning or significance of the following:

Boğazköy	sentence connective	phonetic complement	Sturtevant's Law
Bedřich Hrozný	cuneiform	Akkadogram	mi-conjugation
Indo-Hittite	ductus	determiner	hi-conjugation
h	ideogram	scriptio plena	Wilusa
ergative	syllabogram	syllabary	Çop's Law
clitic chain	Sumerogram		lenition

Exercises

1 Explain briefly the significance of the following forms mentioned in the chapter:

a OHitt. <i>alpa</i>	e Hitt. <i>wātar</i>	i Pal. <i>-āga-</i>
b Hitt. <i>ūk</i>	f Hitt. <i>memiškezzi</i>	j Lyc. <i>rite</i>
c Assyrian <i>išpa(t)ta(l)lu</i>	g HLuv. <i>ziyari</i>	k Lyc. <i>Trqqas</i>
d Hitt. <i>ēšzi</i>	h HLuv. <i>maddu</i>	l Lyc. <i>kbatra-</i>

2 For each of the boldfaced segments or sequences in the following Hittite words, determine whether they would be expected to continue a voiced or voiceless stop

in PIE. Where necessary, indicate if no such determination is possible purely from the spelling.

a	<i>ekan</i> 'ice'	d	<i>pēr</i> 'house'	g	<i>ḥapa-</i> 'river'	j	<i>daluga-</i> 'long'
b	<i>āppa</i> 'back'	e	<i>atta-</i> 'father'	h	<i>daššu-</i> 'mighty'	k	<i>pētan</i> 'place'
c	<i>tūwa-</i> 'far'	f	<i>garāp-</i> 'devour'	i	<i>idālu-</i> 'evil'	l	<i>zakkar</i> 'feces'

- 3 The following are singular active verb forms in Hittite. For each, indicate whether it belongs to the *mi-* or the *ḥi-* conjugation. Hyphens have been added before the inflectional endings.

a	<i>zāi-tti</i> 'you cross'	e	<i>ḥāš-i</i> 'opens'
b	<i>dankueš-zi</i> 'it grows dark'	f	<i>lā-ši</i> 'you unbind'
c	<i>punuš-ta</i> '(s)he asked'	g	<i>pā-u</i> 'let him/her give'
d	<i>paḥṣanu-ddu</i> 'let him/her protect'	h	<i>peḍa-ḥhun</i> 'I took away'

- 4 Convert lines 10–13 in the Hittite text sample (§9.44) into broad transcription. Do not worry about clitic boundaries.
- 5 A Hittite passage (KBo 3.60 li 2–5) reads: *ku-iš iš-tar-ni-iš-mi an-tu-wa-aḥ-ḥi-iš a-ri ša-na-ap az-zi-kán-zi ma-a-an ú-wa-ar-kán-ta-an an-tu-uḥ-ša-an ú-wa-an-zi na-an-kán ku-na-an-zi ša-na-ap a-ta-a-an-zi* 'Whatever man comes among them, they eat him. If they see a fat man, they kill him and they eat him.' Based on your knowledge of Hittite morphology and syntax, determine the Hittite for 'they see', 'they kill', and 'they eat' (two forms for the last one).
- 6 The Anatolian descendants of PIE **deh₃-* 'give', such as Hitt. *dā-*, mean 'take', unlike in the rest of the IE languages. In light of §2.11, provide an explanation for this fact.
- 7 Identify the merisms (§2.43) in the second half of the Luvian text sample (§9.57).
- 8 Name one or more ways that Luvian and Lycian provide important phonological information about PIE.

PIE Vocabulary I: Man, Woman, Kinship

From this chapter forward, a list of roughly a dozen basic roots and lexemes in the proto-language will be given for memorization. Each list will pertain to a particular semantic category. English descendants of the roots are given in SMALL CAPITALS; these are native English or (occasionally) borrowings from Old Norse. When the gloss of a root is at the same time its English descendant, the gloss is given in small capitals. No attempt is made to be exhaustive in listing cognates.

**dhǵhemon-* 'human being': Lat. *homō*, OLith. *žmuō*, OE *guma* (> *bridegroom*)

**h₂ner-* 'man, hero': Ved. *nár-*, Gk. *anēr* (*andr-*), Lat. *Ner-ō* (personal name)

**uiH-ro-* 'man': Ved. *vīrás*, Lat. *uir*, OIr. *fer*, Eng. *WEREWOLF*

Anatolian

**g^henh₂* 'woman': Ved. *gnā*, Gk. *guné*, OCS *žena*, Eng. QUEEN

**ph₂tér-* 'FATHER': Ved. *pítár-*, Gk. *patér*, Lat. *pater*

**máter-* 'MOTHER': Ved. *mātár-*, Lat. *māter*

**suHnú-* 'SON': Ved. *sūnú-*, Lith. *sūnùs*

**dhugh₂ter-* 'DAUGHTER': Ved. *duhitár-*, Gk. *thugátēr*

**bhrāter-* 'BROTHER': Ved. *bhrátar-*, Lat. *frāter*

**syesor-* 'SISTER': Ved. *svāsār-*, Lat. *soror*

**nepot-* 'NEPHEW, grandson': Ved. *nápāt-*, Lat. *nepōs*

10 Indo-Iranian I: Indic

Introduction to Indo-Iranian

10.1. The Indo-Iranian branch consists of two subbranches, Indic and Iranian, and perhaps also a third, Nuristani (see below). Indo-Iranian languages are and have been spoken not only in the present-day countries of India and Iran, but over a wide expanse of Asia from the Black Sea to western China. The branch is bested only by Anatolian as the oldest in the family, with the earliest datable remains dating to the fourteenth century BC. This chapter begins with an overview of Indo-Iranian historical grammar and then covers the development of Indic; Iranian is left for the next chapter.

10.2. The best-known early Indo-Iranian language is Sanskrit. Its oldest known variety is Vedic Sanskrit, preserved in the early collection of hymns known as the Rig Veda and in subsequent literature discussed in more detail below. In this chapter as elsewhere in the book, most cited Sanskrit forms will be taken from Vedic and labeled as such; a form labeled “Sanskrit” is not attested until the later Classical Sanskrit language.

Although Sanskrit has held uncommon esteem throughout the history of IE studies, its linguistic testimony must be balanced against that of its sister Avestan, the oldest preserved Iranian language. Understanding of the Avestan corpus has advanced remarkably in recent decades, and Old or Gathic Avestan, the oldest preserved stage of the language, is structurally very close to Vedic but more archaic in several important respects. The French Indo-Europeanist Émile Benveniste once wrote that “the testimony of Vedic is valuable for its richness, the testimony of Avestan for its fidelity.” The only other older Iranian language preserved in any significant amount is the later Old Persian, the language of royal inscriptions from the Achaemenid Dynasty starting in the late sixth century BC. Old Persian is the only Indo-Iranian language whose remains are securely datable.

10.3. Several languages spoken in a remote region of Afghanistan belong to a group called Nuristani (formerly Kafiri; also sometimes Dardic, a term now properly used of a subbranch of Indic that includes Kashmiri). Many specialists have thought that these languages constitute a separate branch of Indo-Iranian. The fact that they have been poorly studied (they are spoken in inaccessible and war-torn areas) and have no older literature has made their classification difficult. On the one hand, they share several basic sound changes with Iranian, but on the other they preserve some features of reconstructed Proto-Indo-Iranian that are not found elsewhere in the branch.

10.4. It is widely thought that Indo-Iranian forms a subgroup with Greek, Armenian, and Phrygian. The morphological structure of Greek and older Indo-Iranian languages agree in many striking details, and Greek shares numerous similarities with the fragmentary Phrygian as well as the more distant Armenian. However, the issue has not been settled, and for our purposes it is best to regard all four of these as separate branches, even if closely allied at some level.

From PIE to Indo-Iranian

The following is an overview of the major developments from PIE to Proto-Indo-Iranian, that is, the developments that are early enough to be shared by both Indic and Iranian.

Phonology

CONSONANTS

10.5. **Velars.** The defining change to the consonant inventory in Indo-Iranian is its development of the PIE velars. Indo-Iranian is a satem branch (§3.8); it affricated the palatal velars $*k$ $*g$ $*gh$ to $*č$ $*j$ $*jh$, and merged the PIE plain velars $*k$ $*g$ $*gh$ and labiovelars $*k^w$ $*g^w$ $*g^wh$ into just a series of plain velars ($*k$ $*g$ $*gh$). These plain velars, when before a front vowel ($*i$ or $*e$) or the glide $*j$, were then palatalized to the affricates $*č$ $*j$ and $*jh$ (the traditional Indo-Iranian writing of phonetic $[č j j^h]$, pronounced like English *ch* and *j*). This palatalization is often called the Law of Palatals.

To illustrate: PIE $*dek̑m$ ‘ten’ became Indo-Ir. $*dača$ (Ved. *dāsa*, Av. *dasa*); $*ghous-$ ‘hear’ became Indo-Ir. $*ghauš-$ (Ved. *ghóṣas* ‘noise’, Av. *gaoša* ‘ear’); and $*g^wōus$ ‘cow’ became Indo-Ir. $*gāuš$ (Ved. *gāuṣ*, Av. *gāuš*). The Law of Palatals can be illustrated by the development of the first sound in the weak perfect stem $*k^we-k^wr-$ ‘did’ > pre-Indo-Ir. $*ke-kr-$ > $*ce-kr-$ > Indo-Ir. $*ca-kr-$ (Ved. *cakr-*, Av. and OPers. *caxr-*), and in $*k̑ieu-$ > Ved. *cyav-ante* ‘they move’, Av. *šīiauu-āi* ‘I want to drive’.

10.6. **Voiced aspirates.** The voiced aspirates were preserved intact in Indo-Iranian and remain to the present day in Indic, the only IE subbranch still to have them. (In Iranian, the aspiration was lost; see §11.2.) However, if a voiced aspirate was immediately followed by a voiceless unaspirated consonant, a change known as Bartholomae’s Law took place, whereby the aspiration moved to the end of the cluster and the voiceless consonant became voiced. Thus a cluster $*gh̑t$ became $*gdh$, as in Ved. *mugdhá-* ‘dazed’ < $*mugh-tó-$ and Old Avestan *aogadā* ‘he spoke’ (phonetically *aogdā*) < Indo-Ir. $*augdha$ < pre-Indo-Ir. $*augh-to$.

Bartholomae’s Law is particularly interesting in how it affected dental-plus-dental clusters. Recall that dental-plus-dental clusters sprouted a sibilant between the dentals in PIE (§3.36), so $*tt$ > $*tst$ and $*dd$ > $*dzd$. In the case of $*dht$, by a combination of this rule and Bartholomae’s Law, this cluster became $*dzdh$ in Indo-Iranian. In Indic, all these clusters lost the sibilant (so $*tst$ > tt , $*dzd$ > dd , and $*dzdh$ > ddh), whereas in Iranian, they lost the first dental (so $*tst$ > st , $*dzd$ > zd , and $*dzdh$ > $*zdh$ and then ultimately zd). Thus for example Ved. *vittá-* ‘known’ and Av. *vista-* come from $*uid-tó-$; OPers. *azdā* ‘known’ and Ved. *addhá-* ‘surely’ come

from pre-Indo-Ir. **adh-tā*. Note that the name of the Buddha contains an example of Bartholomae's Law (Classical Skt. *buddha*- 'awakened, enlightened').

10.7. Resonants. The two nasals **m* and **n* are preserved intact in all positions: **mṛtijo*- 'mortal' > OPers. *martiya*-, **ne* 'not' > Ved. *ná*, accusative sing. *-*om* > Av. -*am*.

10.8. The fate of the two liquids **r* and **l* is more difficult to ascertain. Iranian and the core of Vedic Sanskrit point to them having merged as **r*: PIE **k^uel-eti* 'turns, moves' > Ved. *cāra*ti, Av. *caraiti*; **bhrāter*- 'brother' > Av. and OPers. *bratar*-. But many varieties of Indic, especially those spoken farther to the east, show *l* as an outcome of both liquids. It may be that the liquids merged as **r* in the western dialect area that would become Iranian and the western dialects of Indic, but merged as **l* in the east, with the original distinction perhaps maintained in the middle. For fuller discussion of the Indic situation, see §10.34 below.

10.9. As for the syllabic resonants, syllabic **r* and **l* merged as **r*: **bhr̥ghent*- 'high, mighty' > Ved. *bṛhānti*, Av. *barazant*-, **u^lk^wos* 'wolf' > Ved. *vṛkas*, Av. *vahrka*-. The syllabic nasals both became **a*: PIE **septm̥* 'seven' > Ved. *saptá*, Av. *hapta*; PIE **n̥* 'not, un-' > Ved. and Av. *a*-. When followed by a laryngeal, the result was a long syllabic resonant: **d̥h₁gho*- 'long' > Indo-Ir. **d̥gha*- (> Ved. *dīrghá*-, Old Av. *darəga*-).

10.10. "Ruki" and creation of **š*. Indo-Iranian created a new sibilant phoneme **š* (**ž* before voiced consonants) from several sources. One source was **s* when preceded by **r* **u* **k* or **i*, a change sometimes known as the "ruki-rule" (compare Balto-Slavic for a similar change, §18.6). This Indo-Iranian **š* became Skt. *ś* and Av. *š*, as in Ved. *tṛṣṇā*, Younger Av. *taršna*- 'thirst' < **tṛs-n*- (root **ters*- 'dry' as in Eng. *thirst*) and Ved. *vakṣayam* 'I cause to grow', Old Av. *vaxša* 'he will grow' (**h₂ueks*-, compare Eng. *wax*). In Iranian, **š* also arose after labial consonants, as in the Avestan nomin. sing. *āfš* 'water' (stem *āp*-).

As expected by the IE voicing assimilation rule (§3.34), the voiced variant **ž* is the outcome before a voiced consonant; this is preserved in Avestan but lost in Sanskrit, e.g. **mizdho*- 'reward' > Av. *mižda*- but Ved. *mīdhām* 'prize' (< pre-Vedic **miždhá*-, cp. Gk. *mīsthós* 'pay'; and see further §10.33).

10.11. The second source of **š* or **ž* is the old palatal velars when they stood before a dental. The voiceless version is illustrated for example by Ved. *našta*-, Av. *našta*- 'died, disappeared' < PIE **nektó*-.

10.12. The third source of **š* was as part of the development of the PIE "thorn clusters" (§3.25), which became in all cases *kš* in Sanskrit but *š*, *xš*, or *ʒš* in Iranian: compare Ved. *kṣétram* 'settlement', *kṣáyati* 'has power', and *kṣáratī* 'flows' with Av. *šōiθra*-, Old Pers. *xšāyaθiya*- 'king', and Younger Av. *ʒzar*- 'flow' (PIE **tkei*-, **tkei*-, **dhg^wher*-).

10.13. Laryngeals. To judge by the metrical evidence from the R̥g Veda and the Gathas in Old Avestan, both consonantal and vocalic reflexes of the laryngeals were still present in Indo-Iranian, though whether the three-fold distinction was still preserved is not known. The vocalic laryngeals became **i*, preserved in Indic but mostly lost in Iranian. See further §§10.36 and 11.7.

VOWELS

10.14. Vowel merger. The defining change in the Indo-Iranian vowel system was the merger of all non-high vowels of PIE, **e* **o* **a*, into a mid to low central vowel that

is written **a*. Thus PIE **bher-ontī* 'they bear' became OPers. *bar-antīy*, **mad-* 'wet' became Ved. *mad-ati* 'is drunk', and **potni_h* 'lady, mistress' became Av. *paθmī*. Analogously, long **ē* **ō* **ā* all fell together as **ā*: PIE aorist **euēghst* 'he conveyed' > Ved. *āvāt*, **māter-* 'mother' > Av. *mātar-*, and **oktō* 'eight' > Ved. *aṣṭā*. The vowel merger only happened after the Law of Palatals and Brugmann's Law (see §10.16 immediately below) had run their course.

10.15. As a result of this merger, the PIE diphthongs **ei* **ai* **oi* all became **ai*, and **eu* **au* **ou* all became **au*. In Sanskrit, **ai* and **au* were monophthongized to *e* and *o*, as discussed further below (§10.38); but they were still diphthongs in the earliest preserved Indic, the fourteenth-century-BC cuneiform documents described in §§10.21ff., and also remained diphthongs in Iranian, as we will see in the next chapter (see §11.8).

10.16. Brugmann's Law. Before the vowel merger noted above, an original **o* in open syllables became lengthened to **ō*, later becoming *ā* by the merger of long vowels. This is known as Brugmann's Law, after its discoverer, the German Indo-Europeanist Karl Brugmann. Thus the 3rd sing. perfect **k^wek^wóre* '(s)he did' (syllabified as **k^we.k^wó.re*) became Ved. *cakāra*, while the 1st singular **k^wek^wór_h₂e* 'I did' (syllabified as **k^we.k^wór.h₂e*) became *cakāra*.

Morphology

10.17. Indo-Iranian has changed the PIE morphological system (as presented in chapters 4–7) only in detail. Thematic and athematic inflection are alive and well in both nouns and verbs, as are all three numbers of singular, dual, and plural. All the tense, mood, and voice categories in the verb, as well as the cases in the noun, are still in use. See the separate discussions in this and the next chapter, where paradigms and other details will also be given.

10.18. In the noun, an important innovation in inflection is the creation of a genitive plural ending **-nām* used with vocalic stems, as Ved. *naḍnām* 'of rivers'. In verbs, the chief innovation was the creation of a passive conjugation with the suffix **-yá-* (a specialization of the PIE accented intransitive suffix **-jé-/ó-*; see §5.29) with middle inflection, as in Ved. *kri-yá-nte* 'they are made', Av. *kiriiēnte* 'they are made' < Indo-Ir. **kr-ǵá-*.

Syntax

10.19. Of great interest for comparative IE syntax is the study of clitic placement in Indo-Iranian, which has shed a great deal of light on Wackernagel's Law, as was discussed in detail in §§8.22–25. Vedic Sanskrit in particular is extremely valuable in this regard because the spelling system quite exactly reflects the operation of sandhi rules (see §10.40 below), whose application is influenced by syntactic movement and constituency, as we have seen in chapter 8 (§§8.31ff.).

Indic (Indo-Aryan)

10.20. Indic (also called Indo-Aryan; see further below) tribes entered India probably during the early to mid-second millennium BC, migrating from the Iranian plateau

northwest of present-day Pakistan into the Punjab in eastern Pakistan, northwest of modern India. One of the hymns of the Rig Veda (1.151) alludes to a legendary journey that may be a distant memory of this migration. The Indus River valley, to the south of the Punjab, had already been the site of an extensive early urban civilization which flourished from c. 2400 to 2000 BC, gradually declining over the next half-millennium. This people left behind short inscriptions in a language that has yet to be deciphered; judging by the material and cultural remains, the Indus Valley Civilization was not Indo-European, but may have been Dravidian. Whether the demise of this civilization was due to the encroaching Indo-Aryans, as used to be thought, or rather to internal or climatic factors, as is generally argued nowadays, is uncertain. The Indus Valley script died with the civilization that had invented it; writing would not return to India until well over a millennium later.

The Mitanni texts

10.21. The earliest Indo-Iranian has been found, of all places, in the Near East. Hittite and Hurrian texts from Anatolia and Syria contain words cited from an early Indo-Iranian language that was spoken by overlords of the Hurrians when they were united into an empire called Mitanni. The Hurrians were a non-Indo-European people probably originating east of the Tigris River who spread westward to become one of the dominant powers in the ancient Near East in the second millennium BC. Starting around 1600 BC their homeland of Hurri was settled by an Indo-Iranian people that was skilled in horse-breeding and chariot-warfare; by perhaps 1500 BC this people had apparently become an elite ruling class among the Hurrians. Under the feudal state that they founded, called Mitanni, the Hurrian lands were united into an empire that lasted until its conquest by the Hittites about 1360 BC.

10.22. From this otherwise unknown language are preserved personal names, divine names, and technical terms pertaining to horse-racing. Names of some numerals are found in compounds referring to laps around a race course, such as *a-i-ka-wa-ar-ta-an-na* '(of) one lap', *pa-an-za-wa-ar-ta-an-na* '(of) five laps', and *na-a-wa-ar-ta-an-na* '(of) nine laps' (compare Ved. *éka* 'one', *pāñca* 'five', *náva* 'nine', and later [post-Vedic] Sanskrit *varṭanam* 'a turning'). In a treaty between Mitanni and the Hittites, the gods *mi-it-ra*, *u-ru-wa-na*, *in-dar*, and *na-ša-at-ti-ya* are called to witness; their names correspond exactly to the Vedic gods Mitra, Varuna, Indra, and the divine twins the Nāsatyas. And in slightly later Babylonian texts are found the color terms *baprumu*, *binkarannu*, and *barittanu* describing horses, which can be equated with Ved. *babhrū*- 'brown', *pingalā*- 'reddish', and *palitā*- 'gray'.

Most of the words could be either Indic or Iranian, but *a-i-ka*- 'one' points to Indic origin, since 'one' in Iranian is **aiwa-* rather than **aika-*. Note also that the Iranian word for 'lap' or 'turn' (Av. *uruaēsa-*) is formed quite differently from the one in the Mitanni speech or Vedic. Various characteristics of the onomastics and divine names also indicate Indic rather than Iranian provenance.

Sanskrit

10.23. The earliest Indic language in which we have significant remains is Sanskrit. This term has both a broad and a narrow sense. Broadly, it refers to any language or dialect belonging to Old Indic, the linguistically oldest preserved stage of Indic;

as it happens, we have only two major dialects belonging to this stage (Vedic and Classical Sanskrit), and they are nearly identical formally. More narrowly, it refers to the somewhat artificial literary language (Classical Sanskrit) discussed below in §10.27, as opposed to the earlier Vedic that we will treat in the next section. We will use the term in its broad sense. ("Old Indic" is especially commonly used in German-speaking lands, where it is rendered in German as *Altindisch*.)

Sanskrit has held a strong grip on the development of IE studies. Even before comparative linguistics came into existence, one can get a feel for the pedestal that Sanskrit would later occupy from Sir William Jones's famous pronouncement in 1786 (quoted in full in §1.14): "The *Sanskrit* language, whatever be its antiquity, is of a wonderful structure: more perfect than the *Greek*, more copious than the *Latin* . . ." As the nascent field of IE studies got its start some decades later, philologists accorded it the greatest importance for the reconstruction of PIE, and early models of PIE differed only in minor detail from Sanskrit itself. The realization, by the 1870s, that Sanskrit too had undergone considerable change (especially in its phonological system), and was not a pristine and unblemished continuation of the parent language, transformed IE studies and solved a number of problems. But far from being "dethroned" thereby, Sanskrit has retained a certain pre-eminence because of, among other things, the age of its oldest texts and the richness and transparency of its morphology.

The Rig Veda

10.24. The earliest preserved Sanskrit is called Vedic, after the Vedas or collection of sacred lore (*véda*- means 'knowledge'). The oldest Veda is the Rig Veda (also spelled Rigveda or Ṛgveda), a collection of 1,028 hymns collected in ten books called *maṇḍalas*. Books II–VII are linguistically the most archaic and are known as the **Family Books**, as each was composed by a particular family of poets. We cannot assign precise dates to the hymns of the Rig Veda; like the Homeric epics, parts of it were composed at different periods and it was transmitted orally over many generations before eventually being committed to writing. But it is reasonable to suppose that the whole collection was completed by the end of the second millennium BC.

The Rig Veda is of paramount importance to IE studies, and it continues to contribute insights into all matters of comparative IE linguistics, poetics, and culture. A difficult text, not all its verses are fully understood. The Vedic poets strove for a deliberately obscure style often densely packed with allusions that are hard for modern readers to recover without careful study of the interconnections among verbal concepts and formulae. (Recall the discussion of one such case in §2.40.)

Other Vedic literature

10.25. The other three Vedas were assembled after the Rig Veda was, and are linguistically younger: the Sāma Veda, Yajur Veda, and Atharva Veda. The verses of the Sāma Veda are drawn almost entirely from the Rig Veda but arranged differently, while the Yajur Veda, divided into the so-called White and Black Yajur Veda, contains not only metrical texts but also explanatory prose commentaries that had accreted onto the transmission of the hymns; this is the earliest preserved Vedic prose. The Atharva Veda contains hymns as well as charms and magical incantations of a more popular and folkloristic type. Exegetical texts like those in the Black Yajur

Veda, as well as ritual directions by the brahmins (priests), developed into separate prose works called Brāhmaṇas. Parallel exegetical traditions developed into other genres of early prose writings – the Āraṇyakas, Upaniṣads, and Sūtras. All these texts are important for understanding the often obscure ritual references in the Vedic hymns themselves, and they contain a wealth of information on early myth and legend that has not yet been fully mined. One sub-genre of the Sūtras, the Gṛhya-Sūtras or ‘household Sūtras’, spawned a set of law texts, the Dharmaśāstras, most importantly the Mānavadharmasāstra or Code of Manu, which contains much ancient legal material of great value for the comparative study of IE law. All these works seem to have been completed by the mid-first millennium BC.

Pāṇinian Grammar

10.26. In the fifth century BC or thereabouts, a grammarian named Pāṇini codified a set of rules for Sanskrit in a work called the Aṣṭādhyāyī. This was the culmination of a long grammatical tradition that is one of the intellectual wonders of the ancient world: it is a highly precise and thorough description of the structure of Sanskrit somewhat resembling modern generative grammar. Roots were set up together with rules for deriving words from them, and the pronunciation of sounds was described in detail. While the analyses are often different from those of modern western linguistics, the work of Pāṇini is very valuable and remained the most advanced linguistic analysis of any kind until the twentieth century.

Classical Sanskrit

10.27. In practical terms, Pāṇini’s codification fossilized the written language; thus was born Classical Sanskrit, a language that became a vehicle for scholarly, religious, and literary discourse, and which is still used today to a certain extent. It achieved a role analogous to that of Latin in Europe during the Middle Ages. Classical Sanskrit is the language of the two surviving Indian epics, the *Mahābhārata* and *Rāmāyaṇa*; the latter was authored by one Vālmiki, while the former, which is roughly eight times the length of the *Iliad* and *Odyssey* combined, was composed and embellished over many centuries. Some of its material is quite ancient and valuable for Indo-European studies. One episode in the *Mahābhārata* has become particularly famous, the philosophical discourse known as the Bhagavad-Gītā, ‘the song of the lord’.

Classical Sanskrit is also the language of the lyric poetry known as *kāvya*, written in a deliberately difficult and ornate style. The greatest master of *kāvya* was Kālidāsa, the author of such poems as the *Meghadūta*. He was also the pre-eminent Classical dramatist; his most famous play, *Śakuntalā*, would later inspire Goethe and other German Romantics. Finally, a gigantic mass of philosophical, scientific, religious, grammatical, mathematical, astronomical, and medical literature is written in Classical Sanskrit, as well as tales and fables such as those of the Hitopadeśa and Pañcatantra, stemming ultimately from the same source as the fables of Aesop.

Classical Sanskrit is not a linear descendant of Vedic, but more like a niece; there are a (very) few formal differences between the paradigms of the two languages. In spite of the great formal similarity, there are many differences in usage and idiom; for example, Classical Sanskrit tended to avoid using finite verbs, favoring instead nominal constructions of various types and non-finite verbal forms.

The terms "Aryan" and "Indo-Aryan"

10.28. As noted above, Indic is also called Indo-Aryan. The term "Aryan" has had a rather complicated history. The Sanskrit word *ārya-*, the source of the English word, was the self-designation of the Vedic Indic people and has a cognate in Iranian **arya-*, where it is also a self-designation. Both the Indic and Iranian terms descend from a form **ārya-* that was used by the Indo-Iranian tribes to refer to themselves. (It is also the source of the country-name *Iran*, from a phrase meaning 'kingdom of the Aryans'.) In the west, various translations of Vedic *ārya-* have been used, most commonly 'nobleman', although we really do not know what its original meaning was. During the nineteenth century, it was proposed that this had been not only the Indo-Iranian tribal self-designation but also the self-designation of the Proto-Indo-Europeans themselves. (This theory has since been abandoned.) "Aryan" then came to be used in scholarship to refer to Indo-European. Some decades later it was further proposed that the PIE homeland had been located in northern Europe (also a theory no longer accepted), leading to speculations that the Proto-Indo-Europeans had been of a Nordic racial type. In this way "Aryan" developed yet another, purely racist meaning, probably the most familiar one today. In Indo-European studies, "Aryan" (and *Arisch* in German) and "Indo-Aryan" are still frequently used in their older senses – "Aryan" to refer to Indo-Iranian (less commonly, Indo-European) and "Indo-Aryan" to refer to Indic.

Sanskrit phonology

Sanskrit historical phonology is very rich and complicated; only important highlights and a few of the more noteworthy details will be discussed here.

GENERAL REMARKS ON THE CONSONANTS

10.29. Sanskrit is the only known older IE language in which the PIE voiced aspirates remain unchanged, though it also made some rather sweeping innovations elsewhere in the system. We find three new series of consonants: voiceless aspirated stops (*ph* *th* etc.); a set of alveo-palatal affricates (usually just called the "palatals," *c* *ch* *j* *jh* *ñ*); and a set of retroflex consonants, dentals pronounced with the tip of the tongue curled backwards and written in Roman transcription with a dot under the letter (*ṭ* *ṭh* *ḍ* *ḍh* *ṇ*). The number of nasals increased to five, one for each of the resultant places of articulation. There was also a symbol called *anusvāra* (indicated in transcription as *ṁ* or *m̐*) and one called *anunāsika* (indicated by *ṁ̐*) that stood for either a reduced nasal sound or nasalization of a preceding vowel. Partly counterbalancing these additions was the satem merger of the PIE labiovelars with the plain velars. The outcome of all these innovations was a neat 25-member system of stops and affricates:

	labial	palatal	retroflex	dental	velar
voiceless	p	c	ṭ	t	k
voiceless aspirated	ph	ch	ṭh	th	kh
voiced	b	j	ḍ	d	g
voiced aspirated	bh	jh	ḍh	dh	gh
nasal	m	ṇ	ṇ	n	ṅ

The voiced aspirates were sometimes reduced to just *h* before vowels, as in the imperative *ihí* 'go!' < **h₁t-dhi*, or the verbal adjective *hitá-* 'placed' from the Sanskrit root *dhā-* 'place'.

The Indo-Iranian velars **k* **g* **gh* (from PIE **k/kʷ*, **g/gʷ*, **gh/gʷh* when not before front vowels) remained unchanged, as did the palatals **ç* and **j* from palatalized **k* and **g*. The affricate **č* (from PIE **ḱ*) became a palatal sibilant transliterated as *ś* (see §10.35 below), while its voiced counterpart **ǵ* (from PIE **ǵ*) fell together with **j* and became *j*. Aspirated **jh* (< palatalized **gh/gʷh*) and **jḥ* (< **ǵh*) fell together as *h*.

GRASSMANN'S LAW

10.30. The Indo-Europeanist Hermann Grassmann discovered that the first in a sequence of two aspirated stops lost its aspiration in Sanskrit. Thus PIE **bʰeudh-etí* 'wakes up' became Ved. *bódhati*, and the participle **bhudh-tó-* became *buddha-* 'awakened' (undergoing Bartholomae's Law also). Grassmann also discovered the same kind of aspiration dissimilation in Greek, though independent of the Indic change (see §12.14).

VOICELESS ASPIRATES

10.31. A number of voiceless aspirates arose out of a combination of voiceless stop plus second laryngeal; for example, **pleth₂-* 'broad' > Ved. *prathi-mán-* 'width', **sth₂-tó-* 'stood' > Ved. *sthi-tá-*. In some seemingly identical phonetic environments, however, aspiration did not occur, as in *pitár-* 'father' < **ph₂tér-*. The details are still insufficiently understood, but most Indo-Europeanists believe the voiceless aspirates to be the result of secondary developments rather than inherited from PIE, as used to be thought (recall §3.6).

DENTAL PLUS DENTAL

10.32. PIE **TsT* and **DzD* from dental plus dental sequences (§3.36) lost the internal sibilant in Indic: Ved. *vyt-tá-* 'turned' < **yrt-tó-*, *vit-tá-* 'found' < **uid-tó-*.

RETROFLEX STOPS

10.33. The retroflex consonants arose under a variety of conditions, most notably from assimilation to a preceding Indo-Iranian **š* or **ž* (§§10.10ff.), as in PIE **oktō* 'eight' > Indo-Ir. **actū* > **aštā* > Ved. *aṣṭā*, and PIE **h₁us-no-* 'burned' > Indo-Ir. **uśná-* > Ved. *uṣṇá-* 'hot'. Since *r* was retroflex, *n* became retroflex *ṇ* after *r* or *ṛ*, as in *dur-nāman-* 'having a bad name', *prṇāti* 'fills'. Note also PIE **nizdos* 'nest' > Indo-Ir. **niždás* > pre-Sanskrit **niždás* > Ved. *niḍás*, where the sound that originally caused the *d* to become retroflex has disappeared (recall §10.10), with compensatory lengthening of the preceding vowel.

Many other words with retroflex stops were borrowed from Dravidian languages to the south. As the Indic tribes moved southward, the number of such loans increased, resulting in a general expansion of retroflexion that even affected originally non-retroflex dentals in native Sanskrit words.

RESONANTS

10.34. As stated earlier, the outcomes of PIE **l* and **r* appear to have varied dialectally. In Sanskrit, both largely merged as **r*. However, forms with *l* are found abundantly, such as *ślōkas* 'poem, type of verse-line' from **kleu-* 'to hear'. But these are rarer in the oldest parts of the Rig Veda, suggesting that they belong to a later infusion of dialectal material from a different part of India, probably the east (Middle Indic inscriptions from the east show a preponderance of *l*). Doublets are not infrequent, as Ved. *riptá-* 'smeared' alongside *liptá-* (< PIE root **leip-* 'to smear, stick').

The syllabic **ṛ* inherited into Sanskrit is often seen rendered in English orthography as *ri*, hence spellings like *Sanskrit*, *Prakrit*, and *Krishna* (Skt. *samskṛtam*, *prākṛtam*, *Kṛṣṇas*). The source of this *ri* is a later development in pronunciation in India.

SIBILANTS

10.35. In addition to the ordinary sibilant **s* inherited from PIE, Sanskrit has two other sibilants of the *sh* variety, a palatal *ś* and a farther-back retroflex *ṣ* (often written *sh* in older handbooks). The first is the ordinary development of PIE **ḱ*, while the second continues Indo-Iranian **š* and represents either PIE **s* with "ruki" (§10.10) or PIE **ḱ* in certain environments.

LARYNGEALS

10.36. The laryngeals were lost in their non-syllabic (consonantal) variants, although not until fairly late; the Rig Veda preserves many words that must scan as though a laryngeal or some remnant of a laryngeal (like a glottal stop) were still present between vowels, a phenomenon called *laryngeal hiatus*. For example, *vātas* 'wind' must sometimes scan trisyllabically as *va'atas*, which comes from earlier **waHatas* (< PIE **h₂ueh₁ntos*). The vocalic laryngeals are continued as the vowel *i* (or sometimes *ī*, perhaps originally only in final syllables): *pītár-* 'father' < **ph₂tér-*; *ábravīt* 'said' < **e-breuH-t*. Traces of word-initial laryngeals are preserved indirectly in the lengthening of preceding vowels in compounds, e.g. *sūnára-* 'mighty' < **h₁su-h₂nero-*, literally 'having good heroes' (**h₁su-* 'good' + **h₂ner-* 'man, hero'). Word-final laryngeals after consonants are preserved as *-i*: *māhi* 'great' = Greek *méga*, both from **megh₂*.

10.37. The "long" syllabic resonants, from original syllabic resonants followed by laryngeals, became *īr* or *ūr* before consonants (e.g. *śīrítá-* 'mixed' < **ḱṛh₂tó-*; *pūrṇá-* 'full' < **pḥ₁nó-*) and *ir* or *ur* before vowels (*tīrátē* 'overcomes' < **tṛh₂-é-*; *purás* 'fort' [genitive sing.] < **pḥ₁H-és*). The conditioning factors determining vowel quality in these outcomes (*i* vs. *u*) are unknown.

GLIDES AND VOWELS

10.38. The glide **j* stayed intact (written *y*), while **u* became a sound transcribed as *v* but still pronounced *w*: *yugám* 'yoke' < **iugom*, *vanóti* 'wins' < **uen-* (cp. Eng. *win*). After the Indo-Iranian merger of PIE **e* **o* and **a*, not much else happened to the vowels except that the Indo-Iranian diphthongs **ai* and **au* monophthongized to *e* and *o*, respectively. Both these vowels were pronounced long and are often

transcribed *ē* and *ō* in older handbooks. The earliest preserved Indic from the Mitanni documents shows these diphthongs still intact: *a-i-ka-* 'one', later Ved. *ēka-*. An example of **au* becoming *o* is the name of the sacred intoxicating drink *sómas* 'soma' from **saumas* (cp. Av. *haoma-*).

The PIE long diphthongs **ēi* **ōi* became Skt. *ai*, and **eu* **ou* became Skt. *au*: Ved. *s*-aorist *dnaiṣīt* '(s)he washed' < PIE **e-nēig^w-s-*; Ved. *o*-stem instr. pl. *-ais* < PIE **-ōis*; Ved. *gáuṣ* 'cow' < PIE **g^wōus*.

ACCENT

10.39. As was noted in §3.32, Vedic preserves the PIE mobile pitch-accent system. The accent markings used in the native script and the Indian grammarians' terms for those markings suggest that syllables preceding the accented syllable had low tone, and that during the pronunciation of the accented syllable the pitch rose, reaching a peak at its end and at the beginning of the next syllable, after which it fell again. The pitch-accent system changed into a stress-accent system in later Sanskrit.

SANDHI

10.40. In colloquial English, the final consonant of a word like *hit* is pronounced differently depending on what sound follows: contrast *hit me* [hɪt̚mij], where the *t* is reduced to a glottal stop, with *hit ya* [hit̚jə], where the *t* merges with the following *y* to form *ch*. As already noted in §8.31, the rules governing such changes in pronunciation at word or morpheme boundaries are called sandhi rules, from Sanskrit *saṁdhī-* 'putting together, transition'. In Sanskrit, these rules are very numerous, and differ depending on whether the boundary occurs inside a word (*internal sandhi*) or between words (*external sandhi*). To give an idea of the phenomenon, consider what happens in external sandhi to word-final *-as*, as in the nomin. sing. of the word *devás* 'god': this form becomes *devá* before most vowels or *s*, *devás* before a *c*, *deváh* before any other voiceless consonant or a pause (the *h* is simply pronounced *h*), and *devó* before a voiced consonant. The Sanskrit sandhi rules are to some extent inherited from common Indo-Iranian, to judge by various traces in Avestan.

External sandhi, especially outside the more artificial language of Classical Sanskrit poetry, did not occur between just any pair of words; its occurrence was governed by syntactic and prosodic conditions that are just beginning to be understood.

Morphology

10.41. Sanskrit has the greatest number of grammatical forms of any ancient Indo-European language. Nouns and pronouns are inflected in eight cases in singular, dual, and plural (the dual in the noun does not have as many case distinctions, however). Personal pronouns have separate full and reduced (enclitic, unstressed) paradigms. Verb tense-stems were still generally built by derivation from roots, as in reconstructed PIE. Verbs in Vedic could be conjugated in six tenses (present, imperfect, aorist, future, perfect, and pluperfect) and in three voices (active, middle, and passive); as in PIE, though, not all verbs were conjugable in all the tenses or voices. There were five moods: indicative, subjunctive, optative, imperative, and (in Classical Sanskrit) precativa, a specialized development of the optative. The augment (Ved. *a-*) could be attached to secondary tenses (imperfect, aorist, pluperfect); when

it was lacking, the form is called an **injunctive** (recall §5.40), which sometimes seems to have had different semantics from augmented forms (e.g. aorist indicative active *āvocam* 'I said' vs. aorist injunctive active *vócam* which sometimes meant 'I will say'; imperfect indicative active *ávinđat* 'he found' vs. imperfect injunctive active *vindat* 'he finds (over and over again)'). The subjunctive in Vedic was less a separate mood and more a simple future tense, while the category that is called the future behaved more like a desiderative or volitional ('I intend to . . .', 'I want to . . .'). The aorist and subjunctive fall into obsolescence in Classical Sanskrit, which relies heavily upon nominal and participial forms to express many concepts that Vedic used finite verb forms to impart. See further §10.47.

VERBS

10.42. Present stem classes. Traditional Sanskrit grammar divides the present stems of verbs into ten classes, which may be exemplified as follows:

I	<i>bhárati</i> 'bears'	VI	<i>tudáti</i> 'beats'
II	<i>ásti</i> 'is'	VII	<i>yunákti</i> 'yokes'
III	<i>dádhati</i> 'puts'	VIII	<i>tanóti</i> 'stretches'
IV	<i>náhyati</i> 'binds'	IX	<i>grbhñáti</i> 'seizes'
V	<i>śṛṇóti</i> 'hears'	X	<i>coráyati</i> 'steals'

From the IE point of view, several of these can be combined. Classes I, IV, VI, and X are all thematic verbs: simple thematics in Class I, verbs in **-je/o-* in Class IV, verbs accented on the suffix in Class VI and usually having zero-grade of the root (called the *tudáti*-class for PIE after the example above; see §5.28), and various derived forms with the suffix *-aya-* in Class X. The other classes are athematic. Class II contains root presents, and Class III reduplicated athematics. Classes VII and IX are nasal-infix presents (§4.18): Class IX contains *set* (laryngeal-final) roots (thus *grbhñáti* 'seizes' < **ghṛbh-né-h₂-ti*, root **ghrebh_h-*), while in Class VII the roots end in some other consonant (so *yunákti* 'yokes' < **ju-ne-g-ti*, root **jeug-*). The nasal present **k_l-né-u-ti* 'hears' (Ved. *śṛṇóti*; root **kleu-*) was the likely model for a whole new class of presents formed by adding **-neu-* to a root; these became the verbs of Class V and most of Class VIII (the latter containing roots that already had an *n* in them from the point of view of the Sanskrit grammarians; thus *tanóti* above was formed from the Sanskrit root *tan*, while in IE terms it was really a **-neu*-present formed from the zero-grade **t_h-*).

The following sample paradigms will illustrate the Vedic active and middle present tense forms of five of the verb classes: *bhárāmi* 'I bear', *dvēṣmi* 'I hate', *śṛṇómi* 'I hear', *yunájami* 'I yoke', and *grbhñāmi* 'I seize'.

	I	II	V	VII	IX
Active					
sg. 1	<i>bhárāmi</i>	<i>dvēṣmi</i>	<i>śṛṇómi</i>	<i>yunájami</i>	<i>grbhñāmi</i>
2	<i>bhāraṣi</i>	<i>dvēkṣi</i>	<i>śṛṇóṣi</i>	<i>yunákṣi</i>	<i>grbhñāṣi</i>
3	<i>bhārati</i>	<i>dvēṣti</i>	<i>śṛṇóti</i>	<i>yunákti</i>	<i>grbhñāti</i>
du. 1	<i>bhārāvas</i>	<i>dviṣvās</i>	<i>śṛṇvās</i>	<i>yuñjvās</i>	<i>grbhñivās</i>
2	<i>bhārathas</i>	<i>dviṣthās</i>	<i>śṛṇuthās</i>	<i>yunikthās</i>	<i>grbhñithās</i>
3	<i>bhāratas</i>	<i>dviṣtās</i>	<i>śṛṇutās</i>	<i>yuniktās</i>	<i>grbhñitās</i>

pl. 1	<i>bhārāmas(i)</i>	<i>dviśmās(i)</i>	<i>śṛṇmās(i)</i>	<i>yuñjmās</i>	<i>grbhñmās(i)</i>
2	<i>bhāratha</i>	<i>dviśthá(na)</i>	<i>śṛṇuthá</i>	<i>yunkthá</i>	<i>grbhñithá(na)</i>
3	<i>bhāranti</i>	<i>dviśānti</i>	<i>śṛṇvānti</i>	<i>yuñjānti</i>	<i>grbhñānti</i>

Middle

sg.

1	<i>bhāre</i>	<i>dviśé</i>	<i>śṛṇvé</i>	<i>yuñjé</i>	<i>grbhñé</i>
2	<i>bhārase</i>	<i>dviśsé</i>	<i>śṛṇusé</i>	<i>yunksé</i>	<i>grbhñisé</i>
3	<i>bhārate</i>	<i>dviśté</i>	<i>śṛṇvé, śṛṇuté</i>	<i>yunkté</i>	<i>grbhñité</i>
du. 1	<i>bhārāvahe</i>	<i>dviśváhe</i>	<i>śṛṇváhe</i>	<i>yuñjváhe</i>	<i>grbhñiváhe</i>
2	<i>bhārethe</i>	<i>dviśáthe</i>	<i>śṛṇváthe</i>	<i>yuñjáthe</i>	<i>grbhñáthe</i>
3	<i>bhārete</i>	<i>dviśáte</i>	<i>śṛṇváte</i>	<i>yuñjáte</i>	<i>grbhñáte</i>
pl. 1	<i>bhārāmahe</i>	<i>dviśmáhe</i>	<i>śṛṇmáhe</i>	<i>yuñjmáhe</i>	<i>grbhñimáhe</i>
2	<i>bhāradhve</i>	<i>dviśdhvé</i>	<i>śṛṇudhvé</i>	<i>yungdhvé</i>	<i>grbhñidhvé</i>
3	<i>bhārante</i>	<i>dviśáte</i>	<i>śṛṇváte, śṛṇviré</i>	<i>yuñjáte</i>	<i>grbhñáte</i>

The forms *śṛṇvé* and *śṛṇviré* are archaic *t*-less 3rd person middles (see §5.15). Vedic and Avestan are both important sources of information on these forms in IE. Note also such forms from Class II as 3rd sing. *śáy-e* 'lies', pl. *śé-re* (from **kéi*-oi and **kéi*-roi), and from Avestan, *mrui-e* 'it is announced', *sōi-re* 'they lie'.

10.43. Aorists. The PIE categories of root, thematic, reduplicated, and *s* (sigmatic) aorist are all found in Sanskrit. A curious and unexplained formation is the 3rd sing. aorist passive in *-i* with old *o*-grade of the root, e.g. *śráu-i* 'he was heard' (**kloy-*). Below are singular and plural paradigms of the root aorist *ásthām* 'I stood' and the sigmatic aorists *ávākṣam* 'I conveyed' and (middle) *ástoṣi* 'I praised'. The forms not attested for these particular roots are in brackets, and forms not attested for any root are left blank.

	Active		Middle	
	root	sigmatic	root	sigmatic
sg. 1	<i>ásthām</i>	[<i>ávākṣam</i>]		<i>ástoṣi</i>
2	<i>ásthās</i>	[<i>ávāṣ</i>]	<i>ásthithās</i>	[<i>ástoṣthās</i>]
3	<i>ásthāt</i>	<i>ávāṣ</i>	<i>ásthita</i>	<i>ástoṣta</i>
pl. 1	<i>ásthāma</i>	[<i>ávākṣma</i>]	[<i>ásthimahi</i>]	[<i>ástoṣmahi</i>]
2	<i>ásthāta</i>	[<i>ávāṣta</i>]		<i>ástoṣhvam</i>
3	<i>ásthur</i>	<i>ávākṣur</i>	<i>ásthiran</i>	<i>ástoṣata</i>

10.44. Other verbal morphology. A few of the other verbal formations in Sanskrit may be mentioned here; this is only a selection. The IE perfect is alive and well and frequently still has stative meaning (§5.49) in Vedic. Two passive verbal adjectives are found, one in *-ná-* and one in *-tá-*, both of good PIE provenance (§5.57). An indeclinable participle known as the gerund grew to be extremely important, especially in the Classical language; it usually indicated action prior to that of the main verb: *ha-tvá* 'having slain', *saṃ-gṛbh-ya* 'gathering, having gathered'. Vedic possessed a large number of infinitives, often with several attested for a single root; by contrast, only one infinitive is found in Classical Sanskrit, in *-tum* (e.g. *ótum* 'to weave'), interestingly just barely attested in the Rig Veda though likely inherited (§5.55).

10.45. An important characteristic of Vedic Sanskrit finite verbs is that they are not marked with accents when they occur in main clauses, unless they are the first word in the clause or verse-line (the *pāda*). This is not to say that such verbs were fully unstressed, as is often stated in the literature. But it does indicate that they were prosodically weaker and probably had lower pitch. The feature is shared with the vocative case of nouns.

NOUNS AND OTHER PARTS OF SPEECH

10.46. The eight cases of the noun were nominative, vocative, accusative, instrumental, dative, ablative, genitive, and locative. All the PIE declensional types are found: consonant-stems, *i*-stems, *u*-stems, long *i*-stems, *a*-stems (from PIE *o*-stems), and *ā*-stems, as well as a few residual heteroclitric *r/n*-stems. Vedic preserves some of the mobile accent paradigms that are reconstructed for PIE, although it has innovated in many details. It also preserves the two different types of feminine *i*-stems, the so-called *devī*- and *vṛkī*-types (after the words for 'goddess' and 'she-wolf', respectively); see §6.71.

Vocatives are accented uniformly on the first syllable, a feature inherited from PIE but only residually present in the other languages. Compare nomin. sing. Ved. *pitā*, Gk. *patēr* 'father' with voc. Ved. *pītar*, Gk. *pāter*.

The following sample paradigms illustrate typical nouns of several classes: *devās* 'god', *śúcis* 'bright', *priyā* 'dear', *adán* 'eating', and *pitā* 'father':

		<i>a</i> -stem	<i>i</i> -stem	<i>ā</i> -stem	<i>nt</i> -stem	<i>r</i> -stem
Sg.	N	<i>devās</i>	<i>śúcis</i>	<i>priyā</i>	<i>adán</i>	<i>pitā</i>
	V	<i>déva</i>	<i>śúce</i>	<i>pṛíye</i>	<i>ádan</i>	<i>pītar</i>
	A	<i>devām</i>	<i>śúcim</i>	<i>priyām</i>	<i>adántam</i>	<i>pitāram</i>
	I	<i>devéna,</i> <i>devā</i>	<i>śúcyā</i>	<i>priyāyā</i>	<i>adatā</i>	<i>pitṛā</i>
	D	<i>devāya</i>	<i>śúcaye</i>	<i>priyāyai</i>	<i>adaté</i>	<i>pitré</i>
	Ab	<i>devāt</i>	<i>śúces</i>	<i>priyāyās</i>	<i>adatás</i>	<i>pitúr</i>
	G	<i>devāsya</i>	<i>śúces</i>	<i>priyāyās</i>	<i>adatás</i>	<i>pitúr</i>
	L	<i>devé</i>	<i>śúcau</i>	<i>priyāyām</i>	<i>adatí</i>	<i>pitāri</i>
Du.	N-A	<i>devā(u)</i>	<i>śúcī</i>	<i>priyé</i>	<i>adántā</i>	<i>pitārā</i>
	V	<i>dévā(u)</i>	<i>śúcī</i>	<i>pṛíye</i>	<i>ádanā</i>	<i>pītarā</i>
	I-D-Ab	<i>devābhyām</i>	<i>śúcibhyām</i>	<i>priyābhyām</i>	<i>adádbhyām</i>	<i>pitṛbhyām</i>
	G-L	<i>devāyos</i>	<i>śúcyos</i>	<i>priyāyos</i>	<i>adatós</i>	<i>pitrós</i>
Pl.	N	<i>devās,</i> <i>devāsas</i>	<i>śúcayas</i>	<i>priyās</i>	<i>adántas</i>	<i>pitāras</i>
	V	<i>dévās(as)</i>	<i>śúcayas</i>	<i>pṛíyās</i>	<i>áantas</i>	<i>pītaras</i>
	A	<i>devān</i>	<i>śúcīn</i>	<i>priyās</i>	<i>adatás</i>	<i>pitṛn</i>
	I	<i>devāis,</i> <i>devébhis</i>	<i>śúcibhis</i>	<i>priyābhis</i>	<i>adádbhis</i>	<i>pitṛbhis</i>
	D-Ab	<i>devébhyas</i>	<i>śúcibhyas</i>	<i>priyābhyas</i>	<i>adádbhyas</i>	<i>pitṛbhyas</i>
	G	<i>devānām</i>	<i>śúcīnām</i>	<i>priyānām</i>	<i>adatām</i>	<i>pitṛnām</i>
	L	<i>devēṣu</i>	<i>śúciṣu</i>	<i>priyāṣu</i>	<i>adātsu</i>	<i>pitṛṣu</i>

That the accusative plurals in *-n* once ended in **-ns* reveals itself in certain sandhi contexts, such as *devāṃs tvām* 'gods you . . .', *nṛṇṃḥ pāti* 'protect men', *sārgāṃr iva* 'like streams' (with *-h* and *-r* from **-s*).

10.47. Compounding in Classical Sanskrit. The more artificial and ornate styles of Classical Sanskrit are famous for the proliferation of lengthy compound nominal forms. Particularly common is the stringing together of words into one great possessive compound (bahuvrihi) modifying some other word in the sentence. Such compounds, of theoretically limitless length, often correspond to whole clauses or sentences in English. A common type is a basically tripartite compound X-Y-Z with the middle member a passive participle, the whole functioning as an adjective meaning 'having (one's) Z Y-ed by X'. In the Buddhacarita (Life of Buddha) by Aśvaghoṣa, for example, a steed is described as *laghuśayyāstaranopagudhapṛṣṭa-* '(whose) back (*pṛṣṭa-*) was covered (*upagūḍha-*) by a short (*laghu-*) bed (*śayyā-*) blanket (*āstarana-*)', and a beautiful woman is compared to a river that is *ṛjusaḍpādapanktijusṭapadmā* '(having) lotuses (*padma-*) that were enjoyed (*juṣṭa-*) by a collection (*pankti-*) of bees (*ṣaḍpāda-*, lit. six-footers) in a row (*ṛju-*)'. In addition to bahuvrihis, the language was fond of compounds expressing essentially a list of things, such as *rogaśokaparitāpabandhanavyasanāni* 'disease (*roga-*), pain (*śoka-*), grief (*paritāpa-*), captivity (*bandhana-*), (and) misfortune (*vyasana-*)'.

Given the Classical authors' penchant for detailed word-painting, these compounds provided a flexible template for extraordinarily rich and dense descriptive passages in both prose and poetry. A consequence of the wide use of such nominalizing, which extended far beyond what we have described here, is a marked decrease in the use of finite verbs. This went hand-in-hand with a simplification of the verbal system that was occurring in the spoken language at the same time; see §10.56 below for more on this.

PRONOUNS

10.48. As noted above, personal pronouns come in two types in Vedic: fully stressed forms in all the cases but the vocative (for example, *tvām* 'thou', *yuvām* 'you two', and *yūyām* 'ye'), and unstressed enclitic forms occurring only in the accusative, dative, and genitive (such as the accusatives *tvā*, dual *vām*, and plural *vas*). The enclitic forms could not be placed at the beginning of a sentence or clause.

10.49. Most of the other pronouns are of IE ancestry as well. Etymologically most obvious are the demonstrative pronoun *sā* (masc.), *sā* (fem.), *tād* (neut.) 'the, this' (exactly cognate with Gk. *ho hē tó*; §7.9); the interrogative pronoun masc. *kás*, fem. *kā*, neut. *kád* (Classical *kim*) 'who, what' from the interrogative stem **kʷo-* (§7.11); and the relative pronoun *yás yā yád* 'who, which, that' (IE **ǵo-*, §7.10).

Syntax

10.50. Vedic syntax is quite similar to that of the other IE languages of comparable date. The many inflections rendered word order fairly free, although recent research is increasingly showing that the freedom was not as great as used to be thought. As in related languages, the beginning of a sentence was a place of prominence; verbs normally come last, but could be fronted for emphasis. Of particular interest is the complex set of rules for the placement of clitics (including both the clitic pronouns and a variety of conjunctions and sentential particles), as discussed in §§8.22ff.

Vedic text sample

10.51. Verses 11–13 of Rig Veda 10.90, the so-called *Puruṣa Hymn*. This hymn is a creation myth, and describes the primeval man, *Puruṣa*, seen as the source of the universe (see §2.32). Verse 12 provides an etiological myth of the traditional caste system of Indian society.

- 11 yāt puruṣam ví ádadhuḥ katidhá ví akalpayan
múkham kim asya káu bahú ká ūrú pádā ucyete
- 12 brāhmaṇò 'sya múkham āsīd bāhú rājanyāḥ kṛtāḥ
ūrú tād asya yād váisyaḥ padbhyām śūdró ajāyata
- 13 candrāmā mánaso jātás cākṣoḥ sūryo ajāyata
múkhaḍ īndraś ca agnís ca prāṇád vāyúr ajāyata
- 11 When they divided up *Puruṣa*, how many pieces did they make him into?
What was his mouth, what were his arms, what his thighs, his feet called?
- 12 His face was the priestly caste, his arms became the princely caste,
his thighs (became) the third caste, from his feet the fourth caste was born.
- 13 The moon was born from his mind, the sun was born from his eyes;
from his face *Indra* and *Agni*, from his breath *Vāyu* was born.

10.51a. Notes. 11. yāt: 'when'. Historically this is the neuter accus. sing. of the relative pronoun, which came to be used as a sort of all-purpose conjunction. ví ádadhuḥ: 'they put apart, divided'; ví 'apart' plus the 3rd pl. imperfect of *dādhami* 'I place' (cp. Gk. *títhēmi*). katidhá: 'into how many parts?' ví akalpayan: 3rd pl. imperfect, 'they made into', from *kalpáyati*, of disputed etymology; it may be a causative that contains the same -p- found in such causatives as *sthā-p-áyati* 'causes to stand' (from *sthā-* 'stand'). kim: 'what?' Notice that interrogative words do not need to be sentence-initial as they normally are in English. asya: 'his', enclitic pronoun, genit. sing. bāhú, ūrú, pádā: 'arms, thighs, feet', all nomin. duals; the lengthening of the final vowels was induced by the PIE dual ending *-h₂ (§6.13). ucyete: 'are called', 3rd dual present passive of the root *vac-* 'call, speak, say', PIE **uek-*, also the root of Gk. (*u*)épos 'word, speech, epic poem' and Lat. *uōx* 'voice' < **uōk-* 's.

12. brāhmaṇò: 'pertaining to the brahmins, priestly, priestly caste'. The word is a vrdhhi-derivative of *brahmān-* 'brahman', itself an amphikinetic (possessive) derivative meaning 'the one of the formulation', from the neuter noun *brāhmaṇ-* 'sacred formulation' (recall §6.29). *Brāhmaṇò* 'sya is the sandhi outcome of underlying *brāhmaṇás asya*. āsīd: 3rd sing. imperfect of *āsti* 'is'. The expected athematic form *ās* (< **e-h₁es-t*) is only marginally preserved in the Rig Veda, replaced nearly everywhere by *āsū*, an innovation formed with the ending -ī. This ending was originally proper only to laryngeal-final roots (< **-H-t*). kṛtāḥ: '(was) made', *tó*-verbal adjective of *kṛ-* 'do, make'. tād . . . yād: literally 'it (was his thighs) that (became) the third caste'. váisyaḥ: 'one who has settled (on the soil), farmer, peasant', a member of the third caste in Indian society; from PIE **ueik-* 'live, dwell', cp. Gk. (*u*)oikos 'home'. ajāyata: 'was born', imperfect of *jan-* 'beget', PIE root **genh₁-*.

13. mánaso: 'from his mind', abl. sing. of *mānas-* 'mind' (Gk. *ménos* 'mental spirit, fighting spirit'). cākṣoḥ: 'from his eyes', ablative dual. agnís: *Agni*, god of fire (cp. Lat. *ignis*, Lith. *ugnis*). vāyúr: *Vāyu*, god of the wind; PIE **b₂ueh₁-iu-*, from **b₂ueh₁-* 'blow', whose *nt*-derivative **b₂ueh₁-nt-o-* gives Ved. *vātas* (read as a trisyllable *va'atas*; §10.36), Lat. *uentus*,

and Eng. *wind*. The second laryngeal at the beginning is directly reflected by Hitt. *hūwant* 'wind' and the Gk. present *á(w)ē-si* 'blows' (=Ved. *vā-ti*).

Middle Indic

10.52. Even as early a text as the Rig Veda is not filled exclusively with archaic Sanskrit forms, but contains (especially in the later books) many that belong to stages of Indic that had undergone additional sound changes. These stages are collectively referred to as **Middle Indic** or **Prakrit**. Throughout the history of Indian literature, Sanskrit and Middle Indic words and texts have coexisted; Middle Indic is thus less a chronological term than one referring to a particular cluster of linguistic developments. The Prakrits, or Middle Indic dialects, are often named after particular regions, but their use spread beyond the purely regional to become characteristic of specific literary genres. For example, in Classical Sanskrit drama different dialects represent the speech of different classes of people: the fairly conservative western dialect Śaurasēnī is used for women and the northeastern dialect Māgadhi for lowborn buffoons. (Compare the use of Doric Greek as the language of the choruses in Greek drama.)

10.53. The term *Prakrit* comes from the Sanskrit word *prākṛtam*, 'made before, original, low, vernacular'; this word was opposed to *saṃskṛtam*, 'put together', hence 'adorned, perfected', already a term for the sacred literary language in Vedic times. The distinction between Classical and Vulgar Latin discussed in chapter 13 is comparable; and just as regional varieties of Vulgar Latin developed into the modern Romance languages, so too did the Prakrits eventually yield the modern Indo-Aryan languages. In literary works, Prakrits are used especially in drama; additionally, a variety known as Ardhamāgadhi is the language of the Jain canon (Jainism is an ascetic philosophy and religion that developed around the sixth century BC).

The Aśokan inscriptions

10.54. Our first written connected texts in an Indic language date are a set of royal rock inscriptions set up in northwest and central India by the emperor Aśoka probably around the mid-third century BC. Aśoka was the third emperor of the Maurya dynasty (c. 325–183 BC), under whose rule India was unified for the first time. The inscriptions are all identical in content but evince regional linguistic differences that are valuable for the comparative study of early Middle Indic dialects. Two inscriptions in the extreme northwest are written in a substantially different dialect, called Gāndhārī.

Pāli

10.55. The immense corpus of canonical and post-canonical works of Theravada Buddhism (an early form of Buddhism now most prominent in Southeast Asia) was written in the variety of Middle Indic known as Pāli. The literary form of this language was fixed relatively early, and by comparison with other Middle Indic dialects it is rather conservative. Tradition says it was the language of the Buddha himself, who came from eastern India, but the features of Pāli are in fact central and western. As Buddhism spread throughout Central and Southeast Asia during the

first millennium AD, Pāli spread with it; the impact of its vocabulary on the languages of this area was significant, and the scripts in which Pāli was written were adopted for writing many languages of Central and Southeast Asia (see further below).

Pāli preserves some archaic features that are drawn from an early Indic dialect different from either Vedic or Classical Sanskrit. For example, the PIE "thorn" clusters all became the voiceless cluster *kš* in Sanskrit, but the ones that were originally voiced clusters sometimes remain voiced in Pāli (as they did also in Iranian), as in Pāli *ug-gharati*, *pag-gharati* 'oozes' < Indo-Ir. **gžharati* (cp. Av. *γžar-* 'flow', and contrast Ved. *kṣārati* 'flows'; the PIE root is **dhgʷher-*). Another archaism is the short *i* in the first syllable of the verb *kiṇāti* 'buys' from PIE **kʷri-néh₂-ti*; the *i* was secondarily lengthened (under the influence of the verbal adjective *krītá-* 'bought') to *krīṇāti* in Sanskrit.

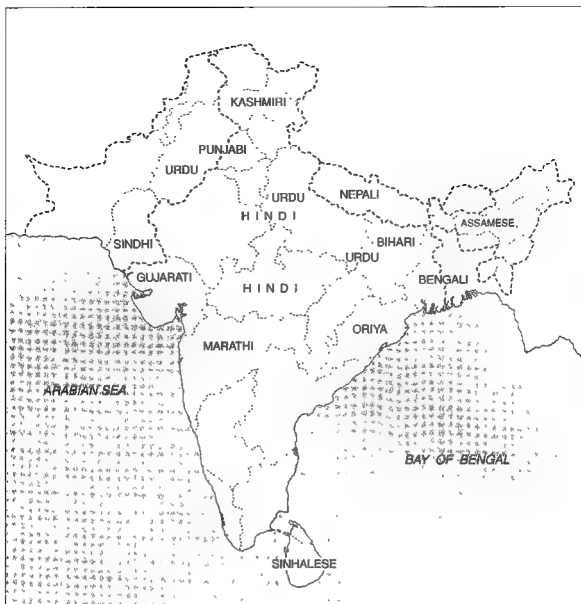
Middle Indic linguistic developments

10.56. Characteristic of Middle Indic historical phonology were the simplification of Sanskrit consonant clusters; the loss of most final consonants; and the weakening or loss of single consonants between vowels. In morphology, both noun and verb inflection was simplified, partly due to the loss of final consonants. Noun cases and stem-classes fell together, and athematic nouns became thematic. In verbs, the several past tenses merged into a single preterite, and the present stem replaced the root as the basic form from which to derive other verbal forms. A new syntactic system known as split ergativity arose, whereby (broadly speaking) the subject of a present-tense verb is in a different case from the subject of a past-tense verb. This came about through the reanalysis of a construction that was already common in Classical Sanskrit, whereby the past passive participle was used to express actions in past time rather than a past-tense finite verb. Thus to say 'Indra slew the serpent', one said literally 'By Indra the serpent (was) slain', with the logical subject (Indra) in the instrumental case. Over time, the logical subject in such constructions became reanalyzed as the grammatical subject. The resulting system is characteristic of the modern Indo-Aryan languages, and was already developing in the later Prakrits.

Development and spread of Indian scripts

10.57. In the earliest inscriptions, two scripts are found, Brāhmī and Kharoṣṭhī. The latter was limited to writing Gāndhārī in northwest India until the fifth century AD, and died out thereafter. Brāhmī was to enjoy much greater fortunes. Its origins have been the subject of controversy, though it is most likely derived, at least in part, from a Semitic alphabet. By around the third century AD, it had evolved into a northern and a southern form. The northern form eventually spawned Devanāgarī, the script of Sanskrit and Hindi. Buddhist missionaries carried related offshoots with them into Central Asia, where they became used for writing Tocharian, Khotanese (a Middle Iranian language), and Tibetan. Knowledge of the script spread as far as Japan, where it greatly influenced the development and organization of the *kana* syllabaries (whose symbols, though, are derived from Chinese).

The southern form, whose signs basically look more rounded, developed into the scripts used for the Dravidian languages in southern India, such as Telugu, Tamil, Kannaḍa, and Malayalam, as well as of the Indic language Sinhalese in Sri Lanka. A variety of the southern type from about the sixth century spread



Map 10.1 Modern Indo-Aryan languages

to Southeast Asia and is the source of the modern scripts of Thailand, Burma, Cambodia, and Laos.

Modern (New) Indo-Aryan

10.58. Between a fifth and a sixth of the world's population speaks a Modern Indic language. A listing of the languages would include well over 200 names, although many of them form large dialect continua without clear divisions. The classification of most of the modern languages is made on the basis of broad geographic areas (zones), including Eastern, Central, Northern, Northwest, and Southern.

The Modern Indo-Aryan languages have continued some of the developments described above for Middle Indic. Diphthongs were often monophthongized and short vowels deleted; final stops and even whole final syllables were often lost. Nominal inflections were reduced in many dialects to just two cases, a direct and an oblique; the number of grammatical genders was reduced to two or, in some areas, none at all.

10.59. The **Eastern Zone** includes Bengali (Bāṅlā), spoken in Bangladesh and northeastern India. Bengali literature is among the oldest of the modern Indo-Aryan languages, starting in the tenth or eleventh century. Linguistically, Bengali is notable for being rather conservative in nominal inflection, with some nouns inflecting in up to six cases; but it has lost grammatical gender. Closely related to Bengali is Assamese, spoken in the state of Assam in eastern India; it also has a relatively rich inflectional system for nouns, and is noteworthy in having lost the retroflex stops. Also belonging to this branch are Bihari (Bihārī) and Oriya (Oriyā).

10.60. The **Central Zone** includes, most famously, **Hindi-Urdu** or **Hindustani**. This represents two literary languages (Hindi and Urdu) of what is essentially a single spoken language. Hindi is spoken across northern India, while Urdu is spoken in northern India and Pakistan (where it is the national language). Literary Hindi is heavily influenced by Sanskrit and written in the same script, while Urdu is rich in Persian and Arabic loanwords and is written in a modified Perso-Arabic script. **Gujarati** (Gujarātī), spoken in Gujarat and the neighboring states of Maharashtra and Rajasthan, preserves all three genders in the noun, unlike Hindi, which only has two (masculine and feminine). **Punjabi** (Panjābī), spoken in the Punjab in northwest India and Pakistan, has developed tones on vowels that originally neighbored voiced aspirated stops; for example, initial voiced aspirated stops have become voiced stops with a low tone on the following vowel. Also belonging to the Central Zone is **Romani** (or **Romany**), the language of the Roma or Gypsies, spoken now primarily in Europe. It apparently separated from the other Central languages around 1000; its modern varieties have borrowed heavily from the languages of the regions where the Roma settled.

10.61. Most prominent of the **Northern Zone** languages is **Nepali** (Nepālī), spoken primarily in Nepal and northeast India.

10.62. The **Northwestern Zone** includes the **Dardic languages**, especially **Kashmiri** (Kashmirī), spoken in northwest India and Pakistan. The Dardic languages share several features with neighboring Iranian languages; Kashmiri, for example, has merged the voiced aspirated stops with the plain voiced stops, probably under Iranian influence. **Sindhi** (Sindhī) is spoken mostly in the Sindh region of Pakistan; it is unusually conservative in retaining final short vowels, but has innovated strikingly in developing a series of implosive stops.

At the other end of the subcontinent is **Sinhalese** (Sinhālā), spoken on the island of Sri Lanka. The ancestors of the Sinhalese came from northern India and colonized the island probably in the fifth century BC. Some inscriptions in an early form of the language, in the Brāhmī alphabet, have been dated to the second century BC; substantial amounts of Buddhist literature were being produced from around AD 1000. Sinhalese merged the aspirated and unaspirated stops.

10.63. Finally, the **Southern Zone** contains **Marathi** (Marāṭhī), spoken in Maharashtra and adjacent states.

For Further Reading

The literature on Sanskrit and Indic is vast; there is, however, no complete comparative grammar, and few major reference works that are either in English or up-to-date. A notable exception is the standard Sanskrit grammar, *Whitney 1896* (still exceptional in spite of its age) and, specifically for Vedic, *Macdonell 1910* (which lists all the forms that occur, by

category) and the smaller (but still very useful) Macdonell 1916. The standard introductory selection of Rig Vedic hymns, with translations, notes, and glossary, is Macdonell 1917. The one comprehensive Sanskrit comparative grammar is in German and incomplete (Wackernagel and Debrunner 1896-); the most recent volume appeared in 1964, and volumes on the verb and syntax are still lacking. The standard edition of the Rig Veda is Aufrecht 1877, and the most commonly used scholarly translation is Geldner 1951-7, unhappily out of print. Readers of French may profitably avail themselves of the 17 volumes of translations of various Vedic hymns and commentary in Renou 1955-69. A new complete translation in English, currently in preparation by the American Indologists Joel Brereton and Stephanie Jamison, will supersede all previous ones when it is finished. Hermann Grassmann's dictionary of the Rig Veda (Grassmann 1873) lists every form and gives the surrounding context for many passages; it is still immensely useful in spite of its often out-of-date definitions. A newly finished and very up-to-date etymological dictionary of both the Vedic and the Classical languages is Mayrhofer 1986-2001. Sanskrit syntax is given fundamental treatment in Delbrück 1886; for more modern analyses of Wackernagel's Law in Vedic and related phenomena, see Hale 1987 (see Bibliography, ch. 8) and subsequent works. A useful new historical and grammatical account of Pāli is Oberlies 2001. The articles of the late Karl Hoffmann on both Indic and Iranian are masterful; they are collected in Hoffmann 1975-92.

For Review

Know the meaning or significance of the following:

Rig Veda	"ruki"	Aryan	sandhi
Nuristani	Brugmann's Law	Indo-Aryan	injunctive
Law of Palatals	Mitanni	retroflex	Middle Indic
Bartholomae's Law	Pāṇini	laryngeal hiatus	Aśoka

Exercises

- 1 Give the outcomes of the following PIE forms in Sanskrit after the operation of Brugmann's Law, changes to the velars, vowel mergers, monophthongization of diphthongs, and changes to the resonants. (Don't forget to apply any relevant pre-Indo-Iranian changes.) Show or describe the intermediate steps. For the purpose of this exercise, both *r and *l became r.

Example: *ǵenos 'tribe, kind, family' became *janas* (*ǵ became j and the two mid vowels became a)

a *ǵ ^w olbhom 'womb'	f *k ^w ormŋ 'doing, action'
b *bhagos 'portion, good fortune'	g *semisk ^w rtom 'thing put together'
c *dhermŋ 'law'	h *nir ^w eh ₂ nom 'blowing out, extinction'
d *ǵhenu 'jaw'	i *ueidos 'knowledge'
e *meghorōǵō 'great king'	j *ieugom 'yoking, joining'

- 2 Vedic has a family of words that includes an adjective *drógha-* 'deceitful' and a root noun *drúh-* 'deceit'.
 - a Which PIE velar or velars are the possible sources for the root-final consonant in this family of words?

- b Part of the paradigm of *drúh-* is accus. sing. *drúham*, nomin. pl. *drúhas*, accus. pl. *druhás*. Are any of these forms not predicted by the operation of regular sound changes?
- c Provide an explanation for any unpredicted forms you identified in b.
- 3 a The neuter of the Classical Skt. interrogative pronoun, *kím* (§10.48), is odd from the point of view of historical Sanskrit phonology. What is odd about it? Can you suggest an explanation?
- b The form *kím* is also odd from the point of view of PIE morphology. Why?
- 4 What would have been the outcome in Sanskrit of the boldfaced sequences in the following PIE forms?
- | | |
|--|--|
| a * h₁elsi 'you go' | e * u₁meuti 'covers' |
| b * lanont- 'setting in motion' | f * h₂usto- 'having shined' |
| c * dikto- 'indicated, pointed' | g * h₂uksont- 'growing' |
| d * u₁sto- 'active' | h * h₂ursto- 'having rained' |
- 5 By Bartholomae's Law and any other relevant sound changes, what would be the outcome in Sanskrit of the following words?
- | | |
|---------------------------------|--|
| a * rudhto- 'obstructed' | c * m₁rdhto- 'neglected' |
| b * labhto- 'taken' | d * dughto- 'milked' |
- 6 A special case involving Bartholomae's Law is the outcome of the cluster *-*gh̥t-*. Based on the following forms, what was the outcome of this cluster? (Ignore the retroflexion of the *ṣ* in *áṣāḍha-*.)
- | | |
|---|---|
| * <i>ṇsegh̥to-</i> > <i>áṣāḍha-</i> 'Invincible' | * <i>u₁gh̥to-</i> > <i>ūḍhā-</i> 'conveyed' |
| * <i>h₃mi₁gh̥to-</i> > <i>mīḍhā-</i> 'urinated' | * <i>d₁ṛgh̥to-</i> > <i>dṛḍhā-</i> 'strong' (<i>r</i> scans long) |
- 7 Apply Grassmann's Law and any other relevant sound changes to the following forms to produce their outcome in Sanskrit:
- | |
|---|
| a * <i>dhedheh₁mi</i> 'I place' |
| b * <i>dhu₁ghroi</i> 'milks' |
| c * <i>bheudhetoi</i> 'wakes up' |
| d * <i>edheg^hhet</i> 'was burning' |
| e * <i>gheghose</i> 'has eaten' (note: apply Grassmann's Law first) |
- 8 As seen in the chart in §10.46, the accusative singular of consonant stems ends in *-am* (*adántam*, *pitáram*). Given Indo-Iranian sound changes, is this expected? If not, say what should have been the outcome, and give an explanation for the form actually occurring. Do not suggest additional sound changes to account for it.
- 9 Identify the class to which the following Sanskrit verbs belong and indicate the type of IE present stem that it continues.
- | |
|---|
| a <i>áste</i> 'sits' (middle) |
| b <i>bhinátti</i> 'splits' (cp. passive <i>bhid₁yáte</i> 'is split') |

- c *pádyate* 'goes' (middle)
- d *nudáti* 'pushes'
- e *pácati* 'cooks'
- f *ṛṇáti* 'fills' (cp. passive *pūryáte* 'is filled')
- g *bíbharti* 'carries' (cp. passive verbal adjective *bhṛtá-* 'carried')
- h *hinóti* 'impels' (cp. 1st pl. root aorist *áhema*)

- 10 One important group of Sanskrit verbs not mentioned in the chapter are the causatives, formed with the suffix *-áya-*. Observe the following pairs of causatives and past participles:

<i>pav-áyati</i> 'causes to cleanse'	<i>pū-tá-</i> 'cleansed'
<i>jar-áyati</i> 'makes old'	<i>jīr-ṇá-</i> 'aged'
<i>mān-áyati</i> 'makes thought (highly of)'	<i>ma-tá-</i> 'thought'
<i>yāj-áyati</i> 'causes to worship'	<i>iṣ-ṭá-</i> 'worshiped'

Given your knowledge of the PIE causative and of sound changes in Indo-Iranian, provide an explanation for the difference in root vocalism (*-a-* vs. *-ā-*) in these causatives.

- 11 a The Vedic text sample in §10.51 contains one verb form showing accent (*ádadhuh*) and four verb forms not showing accent (*akalpayan*, *ucyete*, *āsīd*, *ajāyata*). Explain why some of these are written with an accent and some without.
- b In another Rig Vedic hymn (10.129.2) there is a sentence that reads, *ānīd avātām svadháyā tād ékaṃ* 'That one breathed by its own power, without wind'. Why is the verb *ānīd* 'breathed' accented?

PIE Vocabulary II: Animals

- **eḱuos* 'horse': Ved. *ásva-*, Gk. *híppos*, Lat. *equus*
- **h₂ou-* 'cow': Ved. *gav-*, Gk. *boús*, Lat. *bōs* (*bou-*)
- **h₂ouí-* 'sheep': Lycian *ḫawa-*, Ved. *ávi-*, Lat. *ovis*, Eng. *ewe*
- **k₁uōn* 'dog, hound': Hitt. *kuwan-*, Ved. *śvā*, Gk. *kúōn*, Lat. *canis*
- **suHs* 'pig': Gk. *hūs*, Lat. *sūs*, Eng. *sow*
- **ǵhans-* 'goose': Ved. *harṇsás*, Lat. (*h*)*ānser*, OCS *gōsī*
- **h₂ftkos* 'bear': Hitt. *ḫartaqqaš*, Ved. *fkṣas*, Gk. *árktos*, Lat. *ursus*
- **u₁lk^{*}os* 'wolf': Hitt. *walkuwa-* 'monster' (?), Gk. *lúkos*, Lat. *lupus*
- **mūs* 'mouse': Gk. *mūs*, Lat. *mūs*, OCS *myšī*

11 Indo-Iranian II: Iranian

Introduction

11.1. Languages bearing the designation “Iranian” are by no means limited geographically to Iran. Since ancient times Iranian languages have been spoken over a large section of southwestern and central Asia – from Armenia and Mesopotamia in the west, to the Persian Gulf in the south, all the way into Chinese Turkestan in the east, and well to the north of what is now Iran, Afghanistan, and Tajikistan (the only modern countries whose official languages are Iranian). Chronologically, the Iranian subbranch is divided into **Old Iranian** (until c. 400 BC), **Middle Iranian** (c. 400 BC–c. AD 900), and the modern Iranian languages. By dialect-area it is divided into East and West Iranian, or into Southwest, Central, and Northeast Iranian.

The study of the Iranian languages has made remarkable strides in the last century. The discovery of several Middle Iranian languages in the early 1900s coupled with a steady advance in philological methods in Avestan studies has put Iranian linguistics on an equal footing with Indic. As noted in the previous chapter, Iranian preserves some archaic features that were lost in the former (see the fuller discussion in §11.23 below), but unfortunately the Old Avestan corpus, the most archaic textual material in Iranian, is very small.

Most of this chapter will concern the two Old Iranian languages Avestan and Old Persian, the only ones in which we have texts. We know of several other contemporaneous languages, including Median (§11.27) and Scythian (preserved in some glosses and proper names), spoken in the extreme northwest of Iranian territory. Scythian may have been the Iranian language that came into contact with the early Slavs and from which the latter obtained a few loanwords (see §18.15).

Basic Iranian phonological characteristics

Several features distinguish the Iranian subgroup from Indic; these are innovations that occurred after the Common Indo-Iranian period, or retention of Common Indo-Iranian features that were changed in Indic.

11.2. Deaspiration of voiced aspirates. The voiced aspirates lost their aspiration and became ordinary voiced stops, as in Av. *barai̯nti* and OPers. *bara(n)tiy* ‘they carry’ < Indo-Ir. **bharanti*.

11.3. Spirantization of voiceless stops. Characteristic of Iranian is the development of stops into fricatives in many environments. In the case of the voiceless stops **p* **t* **k*, they became the fricatives *f* *θ* (= [θ] Eng. *th*) and *x* (= [x] German *ch*)

before non-syllabic consonants, as in Av. and OPers. *fra-* 'forth, forward' from Indo-Ir. **pra-*, Av. *caθuuārō* 'four' from Indo-Ir. **caθuāras*; and Av. *xrūra-* 'bloody' from Indo-Ir. **krūra-*.

11.4. Development of the palatals. The Indo-Iranian palatals **ć* and **ǵ(h)* (from PIE **k̑* and **ǵ(h)*, recall §10.5) became Avestan and Median *s* and *z* and Old Persian *θ* and *d*; their intermediate Common Iranian stage may have been the affricates **ts* and **dz*.

11.5. Weakening of **s* to *h*. The old sibilant **s* became weakened to *h* before vowels or resonants in Iranian, as in Av. *hapta* 'seven' (< Indo-Ir. **sapta*) and *ahmi* 'I am' (< **asmi*). This development is similar to that of Greek, as we will see in the next chapter.

11.6. Dental-plus-dental clusters and Bartholomae's Law. As discussed in §10.6, the Indo-Iranian outcomes of the PIE dental-plus-dental clusters, namely **-tst-* (< PIE **-t-t-* and **-d-t-*) and **-dzd(h)-*, lost the initial dental in Iranian and became *-st-* or *-zd-*: Av. *vista-* 'known' (cp. Ved. *vittá-*; Indo-Ir. **uitsta-* < PIE **uid-to-*), OPers. *azdā* 'known' (cp. Ved. *addhá* 'surely'; Indo-Ir. **adzdhā* < earlier **adh-tā*). The progressive voicing assimilation induced by Bartholomae's Law is preserved in Iranian not only in the cluster *-zd-*, but in other clusters as well. In Old Avestan it is still regularly preserved in all contexts, as in *aogadā* (phonetically *aogdā*) 'he said' < **augh-ta* and *aoyzā* 'you said' < **augh-ša*. But in Younger Avestan its effects were largely undone by analogy (for example, 'he said' is *aoxta*).

11.7. Laryngeals. Vocalized laryngeals were lost word-medially in Iranian, as in Av. *duxtar-* 'daughter' (vs. Ved. *duhitār-* < **dhugh₂ter-*), but apparently preserved elsewhere on occasion, as *i* (e.g. in the 1st pl. middle verb ending *-maiōi*, cp. Ved. *-mahi* < **-medhh₂*, and perhaps in Younger Av. *pitā* 'father' < **ph₂tér*, although Old Avestan has *ptā*). The laryngeals in their non-vocalized form left traces in two important ways. First, **h₂* aspirated a preceding voiceless stop, e.g. genit. sing. *paθō* 'path' < PIE **pnt-h₂-es*; and second, Old Avestan preserves so-called laryngeal hiatus (see further §11.21) like Vedic but more faithfully.

11.8. Preservation of diphthongs. In contrast to Sanskrit, Iranian preserved the Indo-Iranian diphthongs **ai* and **au*. These are written as *aē* and *ao* in Avestan, *ai* and *au* in Old Persian: Av. *daēuua-* 'demon', OPers. *daiva-* 'evil god' (vs. Ved. *devás* 'god'); Av. *haoma-*, OPers. *hauma-* 'sacred intoxicating drink' (vs. Ved. *sómas*). Later in Middle Iranian times (§11.35 below), these diphthongs were monophthongized as they had been in Sanskrit.

Avestan

11.9. Avestan is the language of the Avesta, the collection of sacred texts of the Zoroastrian religion. The full Pahlavi (Middle Persian) name, *abestāg u zand*, meaning 'text and commentary', was mistakenly rendered as *Zend-Avesta* in Europe; *Zend* was further misunderstood as a language name, and for well over a century Avestan was known as "Zend." Avestan was introduced to Europe by the French scholar A. H. Anquetil du Perron, who journeyed to India in 1754 to learn about Zoroastrianism from the Parsis (Zoroastrians who had migrated to India in the tenth century) and published the first Avestan texts and translations in 1771. Avestan is classified as East Iranian, and it is believed to have been spoken in an area from the Aral Sea to what is now easternmost Iran.

Zoroastrianism and the Avesta

11.10. Zoroaster or Zarathuštra (Av. *Zarathuštra-*; Zoroaster is the Greek rendering of his name) was a prophet who founded the Mazdayasnian religion (more familiarly called Zoroastrianism). The dates of his life are unknown, although linguistic and comparative evidence points to the late second millennium BC. The Mazdayasnian religion eventually spread throughout pre-Islamic Iran and adjacent territories. Zoroastrians believe in a cosmic dualism in which good spirits (*ahura-*) and evil spirits or demons (*daēuua-*) are in constant conflict, a conflict that will ultimately end with the triumph of good. A supreme being who heads the good spirits is worshiped, Ahura Mazdā or 'wise lord' (Av. *ahurō mazdā*, OPers. *Auramazdā*, ModPers. *Ohrmazd*), whence the term *Mazdayasnian* (*yasna-* means 'worship, sacrifice').

The roots of Zoroastrianism are the inherited Indo-Iranian religion that became Hinduism in India, but Zarathuštra's teachings resulted in many changes, including some that reversed or rejected traditional concepts. A famous example is the Avestan word for 'demon' mentioned above, *daēuua-*, which is the inherited word for 'god' (cp. Ved. *devás*). Nonetheless, many myths in the Rig Veda have analogues in the Avesta, and cognate phrases and formulas are found abundantly in both works (e.g. Ved. *ṛtasya pathā* 'by a path of truth' = OAv. *ašahiia paθō*). Thus the Avesta importantly continues a common inherited religious poetic tradition.

The Avesta consists of several separate texts: the *Yasna* (the liturgy, containing formulaic prayers and hymns); the *Yāsts* (poems, some lengthy, celebrating particular deities and their deeds); the *Vidēvdāt* or *Vendidad* (a legal text concerning punishment and purification); the *Khorde Avesta* (miscellaneous short prayers); the *Nirangestān* (ritual rules and legal matters); and various other texts.

11.11. The linguistically oldest Avestan, called *Old* or *Gathic Avestan*, is confined mostly to the core of the *Yasna*, namely *Yasna* 28–34, 43–51, and 53. These sections constitute the five *Gathas* (or *Gāthās*; *gāθā* 'song') traditionally ascribed to Zarathuštra himself. Also in Old Avestan is the prose ritual text called the *Yasna Haptanḥaiti* or 'Yasna of the seven chapters' (*Yasna* 35.3–41), as well as four sacred formulas preserved in *Yasna* 27 and 54. Old Avestan, as mentioned in the preceding chapter, is grammatically comparable to the language of the Rig Veda. It is therefore usually assumed to be of about the same age, perhaps dating to the late second millennium BC.

11.12. The rest of the Avesta is written in *Younger Avestan* (or *Young Avestan*). As *Younger Avestan* is structurally more archaic than our earliest Old Persian, it is probably several centuries older; a reasonable estimate puts it at the ninth or eighth century BC. It is apparently not a linear descendant of Old Avestan (just as Classical Sanskrit is not a linear descendant but rather a "niece" of Vedic Sanskrit), and the two varieties differ from each other in a few details. For example, all final vowels were long in Old Avestan, regardless of whether they come etymologically from long or short vowels: contrast Old Av. *astī* 'is' with Younger Av. *asti*, both from Indo-Ir. **asti*.

Textual transmission

11.13. Anyone working with Avestan materials needs to be aware of the quality of their transmission, which is sometimes poor. The existing manuscripts are thought to go back ultimately to an edition compiled in Persia during the Sassanian dynasty (AD 224–652). By this point, Avestan had been long dead as a spoken language, and was only used in reciting the sacred religious texts. No direct copy of the

putative Sassanian archetype exists; the surviving manuscripts stem from a much later descendant dating to the eleventh century or thereabouts. The earliest preserved Avestan manuscript dates only to the year 1288, and some of the most important other manuscripts were not produced until the 1600s and 1700s. Clearly, there has been plenty of opportunity for errors to enter into the transmission, both modernizations under the influence of modern speakers and miscopyings.

Script

11.14. An alphabetic script was invented, perhaps in the fourth century, for the express purpose of recording the recitation of the Avestan texts; some of the letters were taken from those of "Book Pahlavi" (see §11.40 below) while the rest were invented afresh. To preserve the recitation accurately, the script was devised to encode much of the phonetic detail of Avestan. (Contrast this with scripts devised by native speakers of a language, which rarely encode such detail because any native speaker can supply the missing information.) Unfortunately the script does not encode all the phonetic information we might like to have; in particular, there is no indication of the placement or nature of the word-stress, for example. The phonetic detail and certain orthographic conventions lend an unfamiliar look to the spelling of many words that are in fact little different from their Sanskrit cognates, like the instr. pl. *daēnuāēibiš* 'by the demons' (Ved. *devébhis* 'by the gods') and the noun *iθiiejah-* 'danger' (Ved. *tyájas-*): *uu* and *ii* represent the glides *u* and *y*; *aē* represents the diphthong *ai*; and an extra *i* before a consonant indicates palatalization of that consonant.

In the manuscripts, words are separated from each other by dots (.), which also regularly appear between the members of a compound (such as *aspō.gar-* 'devouring horses'), and occasionally at or near other morpheme boundaries, such as before an inflectional ending (e.g. OAv. *didrayžō.duiiē* 'you [pl.] wish to hold firmly'). In the standard transcription of Avestan, it is customary to include such dots only when they occur inside a word.

Phonology

CONSONANTS

11.15. **Velars.** Below is a tabular comparison of the Sanskrit and Avestan outcomes of the velars:

PIE	Skt.	Av.
*k, *k ^w	k	k
*g, *g ^w	g	g
*gh, *g ^w h	gh	g
*k, *k ^w before front V	c	c
*g, *g ^w before front V	j	j
*gh, *g ^w h before front V	h	j
*ḱ	ś	s
*ḡ	j	z
*ḡh	h	z

The development of the velars in Avestan is thus fairly straightforward. The Iranian velar stops **k* and **g* (representing PIE **k* **kʷ* **g* **gʷ* **gʰ* **gʷʰ*) and affricates **c* and **j* (representing the same PIE sounds, but before front vowels) remained intact as *k g c j*. The PIE palatals, which had become **č* and **ǵ* in Indo-Iranian, became the sibilants *s* and *z* in Avestan (their outcomes in Old Persian are different; see §11.30), as in *sōire* 'they lie' (Ved. *sére*, PIE **kei-ro-*) and *zairi-* 'yellow, golden' (Ved. *hári-*, PIE **ǵhel-*). Thus Avestan kept three distinct reflexes of palatals, velars before front vowels, and velars elsewhere, in contrast to Sanskrit, which partially merged the first two.

11.16. Spirantization of voiced stops. The spirantization of the voiceless stops has already been dealt with (§11.3). The voiced Indo-Ir. stops *b d g* became fricatives in many environments in YAv., especially word-internally and before voiced consonants of various kinds. These are written *β δ γ*, representing phonetic [v ð ɣ], as in *gərəβnā-* 'seize' (Ved. *grbhñā-*), *daδāiti* 'gives; puts' (Ved. *dādāti*, *dádḥāti*), and *γ(ə)nā* 'woman' (Ved. *gnā-*).

11.17. Iranian **h*. Important for the general look of the language is the further development of Iranian **h* from PIE **s*. Word-internally before *a* it became a sound transliterated as *ḡh*, as in *aḡḡat* 'he will be', the 3rd sing. pres. subjunctive of *ah-* 'be' (cp. Ved. *ásat* 'will be'). A PIE sequence **sʰ* became a sound transliterated *xʷ*, as in *xʷaḡhar-* 'sister' < **sʰesor-*.

VOWELS

11.18. Development of **a*. The development of the vowels in Avestan is complex. We here list some of the changes to affect just one of them, Indo-Ir. **a*. This vowel generally remained intact, but became a sound written *ə* before a nasal, as in the accus. sing. ending *-əm* (Ved. *-am*) and the noun *namō* 'reverence' (Ved. *námas*). The sequence **an-* before a coronal fricative (*s, z, θ*) became *a*, representing a nasalized *a*, as in *maθra-* 'formulation' (Ved. *mántra-*) and *fšuniš* 'cattle-herder' < **pku-jent-s* (from **peku* 'cattle, livestock'; cp. Lat. *pecū* 'cattle'). Final *-as*, most common in the nomin. sing. of thematic masculine nouns and adjectives, became *-ō*, as in *daēuuō* 'demon'; this development also happened before a morpheme boundary in compounds, e.g. *aspō.gar-* 'devouring horses'. A particularly unusual-looking development is the Old Avestan outcome *-əng* from Indo-Iranian **-ans*, as in the accus. pl. *daēuu-əng* 'demons' < PIE **dein-ons*, in the genit. sing. *xʷəng* 'of the sun' < Indo-Ir. **syans* or **suHans*, and in the phrase *dəng paiti-* 'lord of the house' < Indo-Ir. **dans pati-*, PIE **dems potis* (cp. Ved. *dám-patis*, Gk. *des-pótēs* 'master, lord').

11.19. Long **ā* normally remained, but also became *a* before nasals sometimes (e.g. the accus. sing. *daēnam* 'religion' < **-am*). Before **s* it became a sound transliterated as *ā̃*, as in the feminine nomin. pl. *tā̃* 'these' (Ved. *tās*) and the feminine genit. pl. *yāḡham* 'of which' (Ved. *yāsām*).

11.20. Development of **r*. The reflex of syllabic **r* is written *ra*, as in *karatā-* 'done' (Ved. *kr̥tā-*). An interesting development is that an accented **r̥* was devoiced before a following voiceless stop, becoming a sound or sound sequence written *əhr* (e.g. *vəhrka-* 'wolf', cp. Ved. *vṛka-*), and if the stop was a *t*, the combination produced a *sh*-like sound written with a letter transliterated as *š̌*. This happened not only with stressed syllabic **r̥* but also with **-ár-*: *maš̌iia-* 'mortal' (cp. Ved. *mártya-*); *amaša-* 'immortal' (cp. Ved. *amṛta-*).

Note incidentally that these last changes provide important evidence for the existence of a mobile accentual system in early Iranian. A mobile accent is found today in Pashto and a few other modern Iranian languages.

11.21. Laryngeal hiatus. The poetic meter of the Gathas shows that, as in Vedic, certain long vowels count as two vowels (i.e., two syllables) with a hiatus (presumably a glottal stop) between them. The hiatus corresponds to a former laryngeal; we therefore call this phenomenon *laryngeal hiatus*. Thus the nomin.-accus. sing. *dā* 'gift' must scan as two syllables, i.e. *da'ā* (from **da'ō* < **deh₂-os*), whereas the 2nd sing. aorist injunctive *dā* 'you give/gave, you put' does not, as it comes from **deh₃s* or **dheh₃s*, both monosyllables. The poetry of the Gathas is very consistent in preserving laryngeal hiatus where it would be expected on etymological grounds – far more consistent in fact than the Rig Veda, making Old Avestan extremely valuable for the laryngeal theory.

11.22. Sandhi. The complicated external sandhi found in Sanskrit is not present in Avestan. Thus the nominative singular of the word for demon is always *daēnuō*, though its Vedic counterpart ('god') was *devó* only before voiced consonants, and otherwise could be *devá*, *devás*, *devāb*, or *devás*. But etymological final -s's reappear sometimes, especially when a clitic follows, as in *daēnuas.ca* 'and the demon' (cp. Ved. *devás ca*), but sporadically also in other contexts, as in *hauuaitiāsa tanuūō* 'of one's own body' (otherwise *hauuaitiā*). All word-final *t*'s are written with a modified *t* that is transliterated as *ṭ*; this may have been an unreleased *t*. Note, though, the treatment in clitic groups such as *ad-āiš* 'so through them', versus *aṭ* 'so' when alone.

Morphology

11.23. Due to the relative paucity of material, not all inflectional endings are attested; but it is clear that the morphological system of Old and Younger Avestan is essentially the same as that of Vedic Sanskrit, with only some minor differences in detail. Notably though, Avestan, in particular Old Avestan, preserves some archaic features that are not found in Vedic. (Interestingly, most of these features are not found in Younger Avestan either, making the latter in some sense closer to Vedic than to Old Avestan.) For example, in verbs, the distinction between the 1st sing. primary thematic ending **-oh₂* and athematic ending **-mi* is preserved: thematic *paraśā* 'I ask' (versus Vedic *prcchāmi*) but athematic *ahmī* 'I am'. Avestan also preserves traces of proterokinetic inflection in *r/n*-stems, a feature lost in Indic: the word for 'day', *aīiara*, has genit. sing. *aīiṇ* from **aīans*, proterokinetic because of the zero-grade genitive ending **-s*. The word for 'path', nomin. *paṇtā*, accus. *paṇtam*, genit. *paθō* famously preserves the amphikinetic ablaut of its PIE ancestor, **pént-oh₂-s* **pént-oh₂-m* **pnt-h₂-és*, down to the location of the laryngeal in all the forms: the ending *-am* of the accusative scans with laryngeal hiatus, and the laryngeal aspirated the preceding *t* to produce *θ* in the oblique stem, but not in the strong stem where it was separated from the *t* by a vowel.

But retained archaisms are always balanced by innovations; one which may be mentioned is the Younger Avestan extension of the final *-ṭ* of the thematic ablative singular ending *-āṭ* to athematic ablatives (where it replaced the final **-s* used for both the genitive and the ablative), as in *nərəṭ* 'from a man' (vs. genit. *nərəš*).

Old Avestan text sample

11.24. Stanza 4 from Yasna 44, called the *taṭ.θβā.pərəsā Häiti* after its opening words. The poet is asking his patron for a reward for singing his praises and alludes to the fact that a patron who did not remunerate his poet was subject to punishment. Such a reciprocity relation between poet and patron was important in ancient Indo-European cultures; see §2.38.

The translation here is modeled on that of Helmut Humbach.

taṭ θβā pərəsā ərəš mōi vaocā ahurā
 kasnā dərətā zəmcā adš nabāscā
 auuapastōiš kē apō uruuarāscā
 kš vātāi duuənmaibiascā yaogəṭ āsū
 kasnā vañhəuš mazdā dāmiš managhō

This I ask you, tell me straight, Lord:
 Who holds firm both the earth and the heavens
 from falling? Who (holds firm) the waters and plants?
 Who yokes the two swift ones to the wind and the clouds?
 Who (is) the creator of good thought, o Wise One?

11.24a. Notes. *taṭ*: 'this, that', Ved. *tāt*. All the stanzas but the last one in this hymn begin with this line. *θβā*: 'you', accus. sing. enclitic pronoun, Ved. *tvā*. *pərəsā*: 'I ask', 1st sing. present, with *-ā* from PIE **-oh₂* without the addition of *-mi* (vs. Ved. *pyčhāmī*); §11.23. *ərəš*: 'straight', correctly, related to *ərəzu-* 'straight', Ved. *rjū-*. *mōi*: 'me', 1st sing. dat. enclitic pronoun, Ved. *me*; *ōi* and not *āē* was the Avestan outcome of Indo-Ir. **ai* in word-final position and in closed syllables. *vaocā*: 'tell', aorist imperative; Ved. *voca*. This is a reduplicated aorist, PIE **ue-uk-*. Note that the etymologically short final vowel is long here (recall §11.12).

kasnā: 'who', literally 'which man, which person', combination of the masc. nomin. sing. of the interrogative pronoun plus the nomin. sing. of the word for 'man', *nā* (< **h₂nēr*) cliticized and having little meaning. The interrogative has the form *kas-* in close sandhi before the enclitic (see §11.22), and appears alone as *kē* below. *dərətā*: 'holds firm, keeps', 3rd sing. aorist middle injunctive, PIE **dhṛ-to*. *zəmcā*: accus. sing. of 'earth' (= Ved. *kṣām*) plus enclitic conjunction *-cā* 'and' (PIE **kʷe*). The sequence *-cā . . . -cā* means 'both . . . and', as elsewhere in IE. *adš*: 'below', a word occurring only here; related to the adjective *adara-* 'located below', Ved. *ādharā*. *nabāscā*: 'and heavens', accus. pl.; the *-s* only shows up before enclitics.

auuapastōiš: 'from falling', genit. sing. of a *ti*-abstract noun consisting of the prefix *auua-* 'down' plus either **pad-ti-* (from *pad-* 'fall') or **pat-ti-* (from *pat-* 'fly, fall'). *apō*: 'waters', accus. pl., Ved. *apās*. *uruuarāscā*: 'and plants'. The same phrase is in Vedic: *apāśca urvārāśca* 'waters and (plantable) fields' (Atharva Veda 10.10.8). Indo-Iranian **urvarā-* is related to Gk. *āroura* 'plowland' and OIr. *arbor* 'grain', though the details are uncertain.

vātāi: 'wind', dat. sing.; the first syllable scans as two syllables because of laryngeal hiatus (PIE **h₂ueh₂to-*). *duuənmaibiascā*: 'and clouds', dat. pl. of the *n*-stem *duuənman-*. Elsewhere the word has the stem *dunman-*. It is related to Ved. *dhvan-* 'make smoke, form clouds', from PIE **dhyenb₂-*, and perhaps ultimately to PIE **dhub₂-mo-* 'smoke' (> Lat. *fūmus*). *yaogəṭ*: 'yokes', 3rd sing. aorist injunctive. *āsū*: 'the two swift ones', neuter accus. dual, referring to the two horses of a team, or perhaps to two teams. Compare Rig Veda 3.35.4 *yaujmi hāri āsū* 'I yoke the two swift steeds'.

vañhəuš: 'good', genit. sing. of *vañhu-* or *vohu-* 'good' (= Ved. *vāsu-*, from PIE **uesu-*), modifying *managhō* 'thought', genit. of *manō* (= Ved. *mānas*, from PIE **men-os-*, from **men-* 'to think'). 'Good thought' (nomin. *vohū manō*) was a central Zoroastrian concept and one of the manifestations of the Wise Lord, Ahura Mazda. *mazdā*: 'wise', cp. Ved.

medhá- 'wisdom'; both from an Indo-Iranian phrase **m̥s dhā-* 'put (one's) mind'. *damiš-* 'creator', from *dā-* (PIE **dheh₁-*) 'put, create' with the rare suffix *-*mi-* that also appears in the Gk. cognate *thémis* 'right, law' (*'something set down', PIE **dbh₁-mi-*).

Younger Avestan text sample

11.25. From Yašt 14 (the Bahirām Yašt), excerpt from verse 40. An Iranian descendant of the IE dragon-slaying myth (see chapter 2).

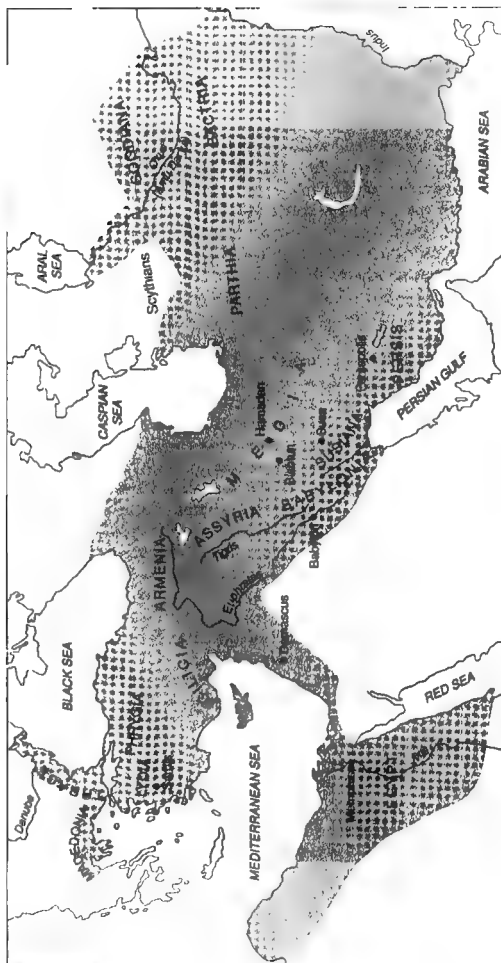
yim θraētaonō taxmō barač
yō janač ažim dahākəm
θrizafanəm θrikamərəðəm
xšuuas.āšim hazarā.yaoxštīm
aš.aojanəm daēuuim drujəm
ayəm gaēθāuuuiō druuantəm

(... the power and the force) which mighty Thraetaona had, who slew the serpent Dahaka, the three-jawed, three-headed, six-eyed, thousand-skilled, the extremely powerful demoniacal monster, evil (and) sacrilegious to living beings...

11.25a. Notes. *θraētaonō*: Thraetaona, a mythical Iranian dragon-slaying hero. *janač ažim*: 'slew the serpent', cognate with Ved. *dhann dhim* in the related Vedic dragon-slaying myth (though the verb form is slightly different, being a thematic rather than an athematic imperfect). The Iranian serpent's name is (Aži) Dahāka. *θrizafanəm*: 'having three jaws', composed of *θri-* 'three' and *zafan-*, from *zafarə* 'jaw', probably an *r/n*-stem, though only the form *zafarə* is attested. *θrikamərəðəm*: 'three-headed'; *kamərəða-* is a special word for the head of a demon. The body-parts of demons had names different from those of corresponding human body-parts. *xšuuas.āšim*: 'having six eyes'; *xšuuas* is 'six', cp. Vedic *ṣát*, PIE **s̥ueks*, and *āši-* is 'eye', cp. Ved. *ākṣi*. *hazarā.yaoxštīm*: 'having a thousand skills'; *hazarā-* 'thousand' is cognate with Ved. *sahāśra-*. *aš.aojanəm*: 'very powerful', containing the prefix *aš-* 'very' and *aojah-* 'strength' (cp. Ved. *ōjas-*). *daēuuim*: 'demoniacal', from Iranian **dayuim*, a *-jo-* adjective formed from **daiya-* (cp. Ved. *devá-* 'god'). *drujəm*: 'monster, lie'; the Iranian term **druj-* or **drug-* refers to all manner of evil things, especially the Lie, the arch-enemy of Truth (Av. *aša-*) in the polarized Zoroastrian world. *gaēθāuuuiō*: 'living beings', dat. pl., with *-āuuuiō* a variant of Younger Av. **-ābiiō* (Ved. *-ābhyas*). *druuantəm*: 'lying, deceptive', a *-*uent-* possessive adjective (§6.40) formed from **drug-* (see above).

Old Persian

11.26. Old Persian is the language of the royal inscriptions of the Achaemenid dynasty of the ancient Persian Empire. The Old Persian inscriptions have the distinction of being the only preserved Old Iranian texts that are authentic originals, written by the very people who spoke the language and free from copyists' errors. The inscriptions come from various sites in western Iran, especially Bisotun (Bisotūn, also spelled Behistūn, Bisitūn), Persepolis, Susa, and Hamadan. Most belong to Darius I (reigned 521–486 BC) and his son Xerxes (reigned 486–465), but inscriptions continue almost to the end of the Achaemenid dynasty, the last ones belonging to Artaxerxes III (reigned 359–338). Over this nearly 200-year period, the language changed quite a bit; the Old Persian of Artaxerxes III had undergone a number of sound changes, including the monophthongization of diphthongs and the erosion of



Map 11.1 The Persian Empire under Darius I

final syllables (and with them, of the case system), that make it look more like Middle Persian than an Old Iranian language. Even the oldest Old Persian is not very archaic compared with Avestan, being typologically quite like late Younger Avestan. This suggests that Younger Avestan and early Old Persian were roughly contemporaneous. Old Persian belongs to the Southwest branch of Iranian.

Median

11.27. Old Persian contains a fair amount of admixture from Median, the language of the earlier Median Empire (c. 700–522 BC) in northwestern Iranian territory. Almost all our knowledge of Median comes from these loanwords, which afford us a partial view of the sound changes that Median underwent and therefore of its filiation within Iranian. Thus the Median loan *aspa* ‘horse’ (native OPers. *asa*) shows that the language patterned with Central Iranian, where IE **k_s* became *sp* (not *s* as in the southwest, where Old Persian was spoken).

The Old Persian script

11.28. Old Persian is written in a cuneiform script that is imitative of Mesopotamian cuneiform, but the actual signs are wholly different. It is a relatively simple syllabary that was apparently designed just for this language. A poorly understood passage of the great Bisotun inscription of Darius I has the king claiming that he devised the script, which could be true. (An inscription purporting to be of the earlier king Cyrus the Great has been shown to postdate his reign.) Old Persian cuneiform was the first cuneiform script to be deciphered; the early breakthroughs were made in 1802 and subsequent years by a school-teacher named Georg Friedrich Grotefend. The existence of bi- and trilingual inscriptions in Old Persian, Akkadian, and Elamite (the language of Persia before it was taken over by the Persians, and subsequently used by them as a court language; it has no known relatives) then paved the way for unlocking the secrets of Mesopotamian cuneiform.

11.29. The script is not well suited for writing Old Persian, and its limitations hamper the analysis of the language’s phonology. Like Mesopotamian cuneiform, the symbols are syllabic, not alphabetic. Only signs for vowels and for sequences of consonant plus vowel (CV) were used. A complete set of CV signs was devised where the vowel was *a*, but where the vowel was *i* or *u* only a subset of the possibilities are represented. Since there were no VC signs, the use of “empty” or “dummy” vowels was frequent, as in the spelling *vi-ša-ta-a-sa-pa-ha-ya-a* for the dat. sing. *Vištāspahayā* ‘for/of Hystaspes’ (the father of Darius I). As this example also shows, long *ā* was represented by adding an extra *a*. Where comparative evidence from Avestan would lead one to expect the presence of *h*, no *h* is actually written before certain sounds; thus the Old Persian equivalent of Avestan *Ahurō Mazda* ‘Ahura Mazda’ is always written *Auramazdā*. Either the *h* was pronounced but not written, or it was simply not there at all. For this reason, one often sees transcriptions like *A^huramazdā* to indicate that etymologically an *h* is expected, but not present in the script. The same is true for *n* at the ends of syllables, as in *antar* (*aⁿtar*) ‘between’. Sometimes the Akkadian and Elamite renderings of Old Persian names provide clues, but the spellings here are often so varied that they are not as helpful as one would wish. No information on the placement or nature of the stress is available; however, some sandhi phenomena are preserved (see below).

Characteristics of Old Persian

PHONOLOGY

11.30. The outcomes of the IE palatals differ considerably from Avestan. In Old Persian, PIE **k̑* and **g̑* became *θ* and *d*, respectively, in contrast with Avestan *s* and *z*: **uik̑*- 'all' > OPers. *viθ*- (Av. *vīs*-), **eǵh₂om* 'I' > OPers. *adam* (Av. *azəm*).

11.31. Also characteristic of Old Persian is the development of the sequence **tr* to *ç* (via an intermediate stage **θr*; *ç* is the standard transcription of a sign in the Old Persian script whose phonetic value has yet to be determined). Compare Vedic *kṣatráṃ* 'kingdom' with Av. *xšaθra*- but OPers. *xšaça*-.

MORPHOLOGY

11.32. Our picture of the morphology of Old Persian has many gaps because of the small corpus, but the forms we do have are in close agreement with Avestan, especially in the verb. The Old Persian noun, however, has undergone more case syncretism (merging of different cases) than Younger Avestan. The dative has been replaced by the genitive, and the ablative has mostly merged with the instrumental and the locative. The relative pronoun had the stem *haya*-/taya- instead of the usual *ya*- in the rest of Indo-Iranian; it also functioned as a kind of definite article and as a particle that was inserted between a noun and a following modifier or appositive, as in *Gaumāta haya maguš* 'Gaumata the priest' (originally meaning *'Gaumata, who [is] a priest'). In Modern Persian, this is continued by the so-called *ezāfe* construction, which utilizes a particle *-(y)e* to connect attributive adjectives and possessives to nouns, as in *mard-e irānī* 'the Iranian man'.

11.33. Other developments include the loss of a distinction between aorist and imperfect; the spread of *a*-stems replacing other stem-types in the noun; and the later contraction of **-iya*- to *-i*-. A development of interest from the viewpoint of Modern Persian is the beginnings of a periphrastic perfect, where the past passive participle was used with a noun in the genitive functioning as the agent, as in *taya manā kartam* 'what (was) done by me' = 'what I have done'. The oblique form *manā* 'by me' eventually became reinterpreted as a subject case, yielding *man* 'I' in Modern Persian. Nowadays one says *man kard-am* 'I have done', where *-am* is a 1st sing. ending added to what was originally a non-finite verb form (a participle; the *-am* in OPers. *kartam* was a neuter sing. ending that was later lost).

Old Persian text sample

11.34. Two excerpts from column I of the great inscription of Darius I inscribed on a steep polished cliff-face at Bisotun, along a caravan route between Baghdad and Tehran. The inscription is bilingual, in Old Persian, Elamite, and Akkadian. The bulk of the inscription details Darius's battles and conquests. The translation is Roland Kent's, with a few modifications. Note that most of the lines end in the middle of a word that then spills over onto the beginning of the next line.

- 1 *adam Dārayavauš xšāyaθiya vazarka xšāyaθiya xšāyaθiy-*
- 2 *ānām xšāyaθiya Pārsaiy xšāyaθiya dahyūnām Višt-*
- 3 *āspahyā puça Aršamahyā napā Haxāmanišiya . . .*

20 . . . θātiy Dārayava-

21 uš xšāyaθiya atar imā dahyāva martiya haya āgriya āha avam u-

22 bartam abaram haya arika āha avam ufrastam aparsam vašnā Auramazdā-

23 ha imā dahyāva tayanā manā dātā apariyāya yaθāšām hacāma aθah-

24 ya avaθā akunavayatā . . .

1-3. I am Darius the great king, king of kings, king in Persia, king of countries, son of Hystaspes, grandson of Arsames, an Achaemenid.

20-24. Darius the king announces: Within these countries, the man who was loyal (?), him I rewarded well; who was evil, him I punished well; by the favor of Ahuramazda these countries that abided by my law, as was said to them by me, thus was it done.

Middle and Modern Iranian

to *ē* and *ō*, the loss of many or most of the tenses, and (particularly in the west) the reduction of the case-system in nouns. Broadly, two groups of Middle Iranian languages can be distinguished, East and West, though the terms are not fully accurate geographically (some East Middle Iranian languages were spoken farther to the west than the West Middle Iranian languages!). This period saw the greatest geographical distribution of Iranian languages, from the northwest coast of the Black Sea to China. The later migratory waves of Turkic peoples into this territory caused the Iranian populations to shrink considerably, but even today most of this area still has at least some pockets of Iranian speakers.

Middle Persian was the only Middle Iranian language about which anything was known until spectacular finds in the twentieth century added several more to the roster: Parthian, Sogdian, Choresmian, Bactrian, Khotanese, and Tumshuqese. Most of these were unearthed in Chinese Turkestan (southwest Xinjiang in western China) – some on the same expeditions that saw the unearthing of the Tocharian languages. Only Middle Persian and Sogdian are known to have close living relatives. Other Middle Iranian languages, such as Sarmatian and Alanic, are known only indirectly and have no preserved literature.

11.36. Most of the Middle Iranian languages were written in scripts derived from the Aramaic alphabet. Aramaic had been one of the official languages of the Persian Empire, and continued to be used in the empire's former territories. Like all Semitic alphabets, it possessed no signs for short vowels and no unambiguous signs for long vowels, rendering it ineffective for recording certain phonetic features of the relevant languages.

West Middle Iranian

11.37. In West Middle Iranian, a stress accent developed on the penultimate or antepenultimate syllable, and final (unaccented) syllables were generally dropped. Nominal inflections were reduced to a system of two cases. The verbal system was more conservative, but the future, aorist, and perfect were lost.

11.38. Parthia was a historical region corresponding to present-day northeastern Iran and neighboring parts of Turkmenistan; in it was spoken Parthian, the official language of the Arsacid dynasty (247 BC–AD 224). The Arsacids themselves were not from Parthia, but when they conquered the Parthians they took over their language. Most of the Parthian we have is preserved in documents from Xinjiang, but there are also some important early royal inscriptions from about 140 BC on. Parthian vocabulary strongly influenced Middle Persian “Book Pahlavi” (see directly below).

11.39. Middle Persian was the official language of the Sassanian dynasty (AD 224–652), but is also known from literature in the ninth and tenth centuries. It was the cult language of Manicheism (or Manichaeism) in Persia, a syncretistic religious movement founded by the Babylonian-born prophet Mani (also Manes or Manichaeus, c. 215–276) and combining the dualism of Zoroastrianism with certain elements of Babylonian religion, Buddhism, and even Christianity. The religion found a wide following in Central Asia and India, as well as for a time in the West (St. Augustine was once a Manichean).

Neither of the two types of Middle Persian discussed below stems directly from the language of the Old Persian Achaemenid inscriptions. Because the literary dialects of Persian, throughout its history, were used for official and administrative purposes,

each stage incorporated forms from whatever other Iranian languages were useful at the time for broadest comprehensibility.

11.40. The language of the Zoroastrian Middle Persian texts or “books” is also called **Pahlavi** (a Middle Iranian descendant of Old Persian *Parθava* ‘Parthian’), a term often applied to Middle Persian as a whole. This “Book Pahlavi” is the standardized written form, using a historicizing orthography developed before Sassanian times. The existence of a Pahlavi translation of the Avesta, with commentary, makes the language valuable for Old Iranian studies. But the script of “Book Pahlavi” is notoriously difficult, since several originally distinct letters have merged, leading to multiple ambiguities in readings. Most of the texts may have been written down in the ninth century from oral traditions, but the surviving manuscripts only date from the fourteenth century, after many corruptions had entered the transmission.

11.41. A much better picture of the pronunciation and vocabulary of Middle Persian comes from the **Manichean Middle Persian** texts, written in a Syriac form of the Aramaic script. It is also a much purer form of the language, as it does not contain the numerous Parthian and Avestan loans that were absorbed into Pahlavi.

East Middle Iranian

11.42. East Middle Iranian was more conservative than West Middle Iranian, as it generally preserved final vocalic endings and, with them, most of the inherited case system in nouns (especially in Khotanese and Sogdian). The verbal system underwent reduction similar to West Middle Iranian.

11.43. The least conservative East Middle Iranian language is **Bactrian**, spoken in the historical region of Bactria between the Oxus River (now the Amu Darya) and the Hindu Kush. Once a remote province of the Persian Empire, it became part of the Hellenistic world after the conquest of Alexander the Great in 331 BC. About 200 years later invaders called the Kushans established an empire there, replacing Greek with Bactrian as the official language but using the Greek alphabet to write it. Bactrian was known for a long time only from coin legends until a 25-line royal inscription was discovered in northeastern Afghanistan and published in 1951. Recent discoveries have added hundreds of well-preserved legal documents to the roster, as well as additional inscriptions.

11.44. Two languages, **Khotanese** and **Tumshuqese**, are known from finds in Chinese Turkestan and are collectively referred to as **Saka**. Remains of Tumshuqese were discovered in several villages near Kucha, and while our knowledge of it is slim, it is clearly more archaic than Khotanese. The latter is known from much larger finds; it was spoken in the medieval kingdom of Khotan (centered around the present-day city of Hetian on the southern Silk Road). The Khotanese documents include Buddhist texts, letters, and legal documents and date from the seventh to the tenth centuries. Both Khotanese and Tumshuqese are written in varieties of the Brāhmī alphabet (see §10.57). Neither language has any living descendants, but Wakhi, spoken today a bit to the west in the Wakhan corridor of northeast Afghanistan, is closely allied (see §11.51 below). Khotanese preserves six cases in the noun and several noun classes; it has also preserved all the moods in the verb, but lost most of the tenses.

11.45. **Sogdia** or **Sogdiana**, where **Sogdian** was spoken, is an ancient land in Central Asia already mentioned in the Old Persian inscriptions. Its capital was Samarkand (now in Uzbekistan), which became an important node on the Silk Road. Sogdian

merchant colonies were established to the east, and Sogdian practically became a trade language of Chinese Turkestan. It is preserved in Manichean, Christian, and above all Buddhist texts, especially from the Turfan Oasis and Dunhuang deep inside Xinjiang and Gansu provinces, respectively. It is also preserved in non-religious writings. Like Khotanese, it is rather conservative in retaining most nominal cases. It is particularly striking in having innovated a very rich verb system. A variety of Sogdian different from the preserved variety was ancestral to modern Yaghnobi (see §11.51 below).

11.46. The inhabitants of the medieval land of Choresm (Khwarezm), located along the Amu Darya south of the Aral Sea along the present-day Turkmenistan-Uzbekistan border, spoke **Choresmian** (also Chorasman or Khwarezmian). The language is known mostly from an interlinear translation of a medieval Arabic encyclopedia written in 1135, although scattered writings from much earlier have been found on a variety of materials. Linguistically it is similar to Sogdian, but has undergone further development.

Modern Iranian

11.47. The modern Iranian languages share the syntactic phenomenon known as split ergativity, wherein the subject of a past-tense verb is in an oblique case, while the object appears in the nominative case (normally used for the subject). This is a development from an Old Iranian stylistic feature (shared also by Sanskrit and likewise continuing as split ergativity in many modern Indo-Aryan languages) whereby past participles were used instead of finite past-tense verbs to express actions in past time (cp. §10.56).

Modern Iranian languages are spoken over a wide area, from the Caucasus in the northwest (where Ossetic is spoken) to northern Oman in the south (where Kumzārī is spoken) to Xinjiang in the east (where Sarikoli is spoken). We will only mention the more prominent Modern Iranian languages below.

11.48. Modern West Iranian. By far the most familiar language is **Modern Persian** or Farsi (Fārsī), the national language of Iran and one of the two national languages of Afghanistan. The variety spoken in Afghanistan is also called Dari. A dialect of Persian known as **Tajiki** (older Tadjhiki) is spoken in Tajikistan and neighboring countries. Two other modern West Iranian languages deserve mention, **Kurdish** (Kurdi) and **Balochi** or **Baluchi** (Balōči). Kurdish is spoken in Iraq, Iran, Turkey, and Russia; Balochi is spoken primarily in Balochistan (Baluchistan) in southeast Iran and southwest Pakistan.

11.49. Modern East Iranian. The other national language of Afghanistan, **Pashto** (Paštō, also spelled Pushtu in older literature), is the most populous East Iranian language and, aside from Persian, the modern Iranian language with the longest and finest literary tradition (from the sixteenth century). Pashto is the only modern Iranian language to have accent oppositions in morphological classes. This system very indirectly continues the mobile accent of older Iranian.

11.50. Ossetic, with half a million speakers in Turkey and Georgia, is of interest because of its isolation from other Iranian languages over the last two millennia; it is the last remnant of the essentially unknown Middle Iranian dialect area that included Sarmatian, and is said to descend from Alanic. It has taken on many features of the neighboring Caucasian languages, such as glottalized consonants and a system of nine cases in the noun.



Map 11.2 Selected Modern Iranian languages

11.51. Two other Modern East Iranian languages, **Yaghnobi** (Yaghnōbī), spoken by a small population in Tajikistan, and **Wakhi** (Wakhī), spoken in the Pamir and far western China, deserve mention for linguistic reasons. Yaghnobi is surely descended from an unknown (non-literary) variety of Sogdian, and is the only modern Iranian language that still uses the augment to mark past-tense verb forms. Wakhi is important for preserving some archaic features not even found in Avestan, such as past participles in *-n-* that go back to the PIE suffix **-nó-* (§5.57), not otherwise found as a participial suffix in Iranian.

For Further Reading

All the Iranian languages, literatures, and cultures, past and present, are treated exhaustively in the ongoing *Encyclopaedia Iranica* (Yarshater 1982–), although many years will pass before it is completed. The languages are given very able and up-to-date historical treatment in the articles in Schmitt 1989. The only complete dictionary of Avestan is Bartholomae 1904, which also contains Old Persian; it is now to some extent out of date. No Iranian etymological dictionary has yet been written, but in Manfred Mayrhofer's etymological dictionary of Sanskrit (see the previous chapter) can be found all the Old Iranian material that has Sanskrit cognates (if one knows which Sanskrit word to look under). The one complete grammar of Avestan is also an older work, Reichelt 1909. The standard critical edition of the Avesta is Geldner 1886–96, whose introduction contains an important philological evaluation of the manuscripts and their relationships to one another. A superb historical grammar of Avestan is Hoffmann and Forssman 1996. Avestan morphology has been the subject of two recent excellent studies in French, Kellens 1974 (root nouns) and Kellens 1984 (the verb). The difficulty of the Gathas has led to many differing interpretations, and several translations with commentary are available, including Insler 1975 and Humbach 1991. The texts, grammar, and vocabulary of Old Persian are easily accessible in Kent 1953, though a few additional texts have been discovered since. The corpus of Iranian inscriptions has been appearing in the

ongoing *Corpus Inscriptionum Iranicarum* (London, 1955–). Many excellent articles on Iranian by Karl Hoffmann can be found in Hoffmann 1975–92 (see Bibliography, ch. 10).

For Review

Know the meaning or significance of the following:

Old Iranian	Zoroastrianism	Achaemenid
Middle Iranian	Gathas	Georg Friedrich Grotefend
Avesta	laryngeal hiatus	Pahlavi

Exercises

- 1 Using the table in §11.15, for each of the following Sanskrit–Avestan cognate sets determine which PIE velar underlies the boldfaced segments. Some forms may have more than one answer.

- a Av. *kainiā-*, Ved. *kanyā* 'unmarried girl'
- b Av. *razišta-*, Ved. *rājiṣṭha-* 'most upright'
- c Av. *zaranīa-*, Ved. *hiranya-* 'golden'
- d Av. *zanga-* 'ankle', Ved. *jāṅghā-* 'lower leg' (two different velars)
- e Av. *karša-* 'furrow', Ved. *kṛṣāti* 'cuts a furrow'
- f Av. *hacaite*, Ved. *sacate* 'accompanies'
- g Av. *dādaresa*, Ved. *dadārśa* 'I saw'
- h Av. *aojō*, Ved. *ójas* 'power'
- i Av. *jaiṇti*, Ved. *hānti* 'smites'
- j Av. *genā*, Ved. *gnā-* 'woman'
- k Av. *garēma-*, Ved. *gharmám* 'heat'
- l Av. *-maēzaiti*, Ved. *méhati* 'urinates'
- m Av. *adrujīant-* 'not deceitful', Ved. *drúhyati* 'deceives'
- n Av. *-zānaiti* 'recognizes', Ved. *jānāti* 'knows'
- o Av. *sāstī*, Ved. *śāste* 'teaches'
- p Av. *cisti-*, Ved. *cittis* 'thought'

- 2 Below are given various Vedic words. Predict what the Avestan cognates would look like based on the sound changes enumerated in §§11.2–8 and 15–20.

- a *spṛdhi-* 'eagerness'
- b *trātā* 'protector'
- c *śróta* 'you (pl.) hear'
- d *bhāram* 'I carried'
- e *mātāras* 'mothers' (nomin. pl.)
- f *dhenú-* 'female'
- g *tvām* 'you' (accus. sing.)
- h *māsam* 'month' (accus. sing.)

- 3 The "Hoffmann" suffix (§6.35) was originally identified on the basis of the Old Avestan word *māθrā* (nomin. sing.), stem *māθrān-* 'possessing a *māθra-* or religious formulation (mantra)'. The *ā* in the nominative and accusative scans in the Gathas as two syllables. Explain how this fact allows us to deduce the shape of the suffix.

- 4 Avestan has the form *kəhrpəm* 'body', accusative singular of a root noun whose cognate in Vedic only occurs in the instrumental singular, *kpá*. Given the rules in §11.20, what would you predict the Vedic accusative singular to have been?
- 5 Given §10.37 and the forms below, what was the Avestan outcome of syllabic resonant plus laryngeal before vowel?
- a Av. *tarō*, Ved. *tírās* 'through' c Av. *parō*, Ved. *purás* 'in front'
- b Av. *garō*, Ved. *gírās* 'praises' (noun) d Av. *-sparaŋ*, Ved. *sphurat* 'hurried'
- 6 a Given that Ved. *khā-* 'well' is cognate with Av. *xā-*, and that Ved. *pathás* 'from the path' is cognate with Av. *paθō*, and finally that Ved. *śaphá-* 'hoof' is cognate with Av. *safa-*, what were the Avestan outcomes of the Indo-Iranian voiceless aspirated stops?
- b What would you reconstruct as the Indo-Iranian ancestor of Ved. *kumbhá-* 'pot' and Av. *xumba-*? Explain.
- c The word for 'path' in Vedic has nomin. sing. *pánthās*, accus. sing. *pánthām*, and weak stem *path-*. Contrast this with its Avestan cognate, discussed in §11.23. How do you account for the *-th-* in the Vedic nominative and accusative forms?
- 7 Comment on each of the following morphological differences between Vedic and Avestan:
- a athematic ablative singular: Ved. *mánasas*, Av. *manarhaŋ*
- b perfect participle: Ved. *vidvān*, *vidvāṁsam*, *viduṣ-*, Av. *viduua*, *viduuaŋhem*, *viduṣ-*
- c *r/n*-stem nominative sing.: Ved. *yákṛt* 'liver', *śákṛt* 'excrement'; Av. *baēuuare* 'myriad', *aliare* 'day' (ignore the schwa)
- d *r/n*-stem instrumental pl.: Ved. *udá-bhis* 'with waters' (from **udh-*), Av. *baēuuare-biš* 'with myriads' (ignore differences in the endings)
- 8 a The Old Avestan 2nd sing. imperative of the verb 'be' is *zdī*. What is its history?
- b A supposed sound law of PIE date was Siebs's Law, whereby a voiced aspirate became unvoiced following **s*. Comment on the validity of this law in light of forms like *zdī* and *mižda-* 'reward' (§10.10).
- 9 Given the remarks in §11.29 and the example there, render the first three lines of the Old Persian selection in §11.34 as you believe they would have been written in the Old Persian syllabary. Use hyphens to separate signs in the syllabary. Remember there are no signs representing consonants standing alone or vowel-consonant sequences. Assume that where a *Ci* or *Cu* sign is needed, it exists.

PIE Vocabulary III: Food and Agriculture

**aǵros* 'field': Gk. *agrós*, Lat. *ager* (*agr-*), Eng. *acre*

**h₂erh₃-* 'plow': Gk. *aróō* 'I plow', Lat. *arāre* 'to plow'

**iugom* 'yoke': Ved. *yugám*, Gk. *zugón*, Lat. *iugum*

**seh*₁- 'sow (seed)': Lat. *sē-men* 'seed', Lith. *sėjū* 'I sow'

**ǵr̥h₂-no-* 'grain': Lat. *grānum*, Eng. CORN

**h₂melǵ-* 'MILK': Gk. *amélgō* 'I milk', Lat. *mulgēre* 'to milk'

**melit-* 'honey': Hitt. *milīt*, Gk. *méli* (*mélit-*), Lat. *mel*, Eng. MILDew

**medhu-* 'sweet drink': Luv. *mattu-* 'wine', Ved. *mádhu-* 'honey', Gk. *méthu* 'wine',
Eng. MEAD

**sal-* 'SALT': Gk. *háls*, Lat. *sāl* (*sal-*), OCS *solī*

**melh₂-* 'grind': Hitt. *mallai* 'grinds', Gk. *múlō* 'mill', Lat. *molō* 'I grind', Eng. MEAL

12 Greek

Introduction

12.1. It is a remarkable fact that Greek has remained Greek over its 3300-year written history. Throughout this time, its dialects never developed into mutually incomprehensible languages, and are at all times recognizably woven out of the same linguistic fabric.

We do not know when Greek (or pre-Greek) speakers first came to Greece nor what languages were spoken there before. Much has been written about the pre-Greek linguistic situation, most of it speculative and inconclusive. For example, it is usually claimed that a number of words containing an element *-nth-* come from a pre-Greek substrate language, such as *Kórinthos* ‘Corinth’, *labúrinthos* ‘labyrinth’, and *asáminthos* ‘bathtub’; but we can say nothing conclusive about such words or the language(s) they may have come from.

By the time we first meet Greek in written documents, in the second half of the second millennium BC, Greece and parts of the western coast of Asia Minor were already Greek-speaking. The earliest preserved Greek is written in Mycenaean (or Mycenaean), the official dialect of the Mycenaean civilization. Mycenaean inscriptions are preserved on clay tablets and ceramic vessels found on the isle of Crete (primarily at Knossos) and the mainland (primarily at Pylos), always from Mycenaean royal cities. The dating of the Mycenaean tablets has been in flux for some decades, but it appears that the island inscriptions are older than the mainland ones, the oldest being the so-called Room of the Chariot Tablets from Knossos (c. 1400–1350 BC); the Pylos tablets date to c. 1200 BC. Mycenaean is most closely related to Arcado-Cypriot (see below on the classification of the Greek dialects).

12.2. Mycenaean is written in a syllabic script called Linear B, whose decipherment was announced by Michael Ventris in 1952. It was developed from the earlier Linear A script, which was used to write the language of the Minoan civilization but remains undeciphered. The Linear B syllabary is poorly suited to writing Greek. It contains only signs representing single vowels or open syllables (consonant–vowel), does not distinguish between short and long vowels, between *r* and *l*, or usually between voiced, voiceless, or aspirated stops. Consonants at the ends of syllables are rarely written. Because most of the texts are inventories (lists of offerings, to cite a common type), our knowledge about the structure of this dialect is limited. Nonetheless, what we do have is extremely valuable, as the dialect is a full 500 years older than the next-oldest inscriptions of any length. Mycenaean is famous for preserving the inherited labiovelars as a distinct series of stops (transliterated with

the letter *q*), unlike the later dialects (see §12.16), and it has many inflections, words, and phrases found otherwise only in the Homeric epics.

12.3. The period from the demise of Mycenaean civilization to the earliest appearance of alphabetic Greek in the eighth century BC has traditionally been called the "Greek Dark Ages." During this time Greece plunged largely into illiteracy and stagnation; there was a drastic drop in population and much movement of refugees, and a unified Greek civilization no longer existed. Nonetheless, at least one inscription from the period is known, a skewer found in Cyprus with a name incised in the Cypriot syllabary (see below) and dated to about 1050 BC.

The cultural regionalism that began in the "Dark Ages" persisted for many centuries, and in no way is this better reflected than linguistically: the history of post-Mycenaean Greek is essentially the history of its local dialects, even well into the Classical period (480–323 BC). Not until the following age, the Hellenistic period, would Greece again become both culturally and linguistically unified.

12.4. The period from the appearance of the first alphabetic inscriptions, in the early eighth century BC, to 480 BC is termed the Archaic period. Near the beginning of this time is probably when the earliest known Greek literary works – the two surviving Greek epics, the *Iliad* and the *Odyssey* – reached their present form. The language of these poems, called Epic (or Homeric), is a literary mixture from a variety of different periods. Representing the culmination of a long oral-poetic tradition stretching back into the dim past, the epics were not composed by a single person; the bard "Homer" was as remote and shadowy to the ancients as he is to us, and may well be a fiction. Rather, they were elaborated regionally by generations of illiterate poets until they were finally committed to writing perhaps around 700 BC. Aside from their status as two of the greatest works of literature ever composed in the Indo-European-speaking world, the *Iliad* and the *Odyssey* are of supreme importance for Indo-Europeanists because of their seemingly inexhaustible supply of archaic forms, phrases, and cultural material. Some phrases and turns of language in the epics date to before Mycenaean times, and scattered throughout are some inherited poetic formulae from PIE. We will discuss the study of these works in more detail in §§12.58ff.

Epic is also the language of the thirty-three so-called Homeric hymns, each addressed to a specific deity, and of the works of the first Greek author whose name we know, Hesiod (late eighth century BC). Some of the Homeric hymns contain important ancient material, but all of them were composed much later than the two epics.

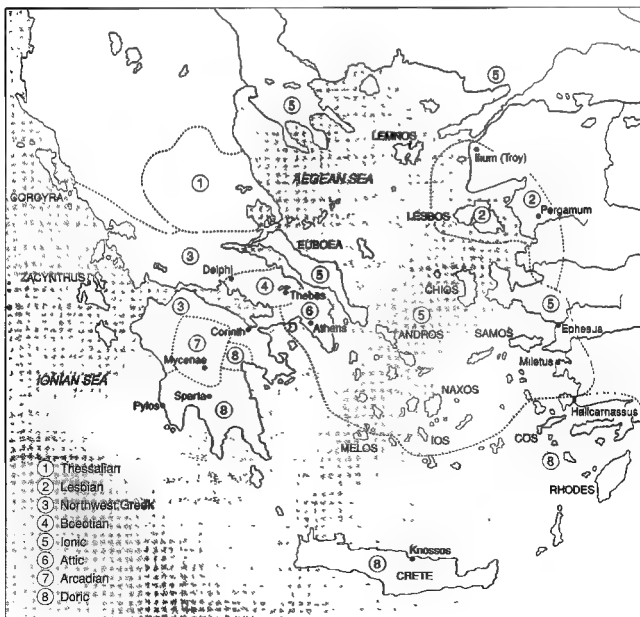
Epic has at its base the dialect called Ionic, the variety of Greek spoken in Ionia (the coast of western Asia Minor), the island of Euboea, and the Cyclades islands in the Aegean Sea. Epic also contained, mixed in with Ionic, a substrate of Aeolic forms. Aeolic was a group of dialects spoken on the island of Lesbos and over the mainland Greek areas of Thessaly and Boeotia. Particularly prominent in other literature from the Archaic period is the Aeolic dialect of Lesbos, called Lesbian, in which was written the lyric poetry of such figures as Sappho, Alcaeus, and Stesichorus from the seventh and sixth centuries BC.

The Classical period

12.5. Due to the cultural and political hegemony enjoyed by Athens in the fifth century BC, it was Attic, the dialect of Athens, that became the standard literary

language during the **Classical period** (480–323 BC). Attic was closely related to Ionic, and the two form a group called **Attic-Ionic**. Early on Attic became the language of drama (essentially an Athenian creation), as represented by the tragedies of Aeschylus, Sophocles, and Euripides, and the comedies of Aristophanes. Soon it became used for prose as well, by historians like Thucydides and Xenophon; philosophers like Plato, Aristotle, Theophrastus, and Epicurus; orators like Isocrates and Demosthenes; and hundreds of other authors. But the earliest Greek prose, that of philosophers such as Heraclitus, was written in Ionic.

12.6. Attic-Ionic, the Aeolic dialects, and Mycenaean are all part of the **East Greek** dialect group, which includes also the non-literary dialects **Arcado-Cypriot** (or **Arcado-Cyprian**, consisting of Arcadian, spoken in the central Greek region of Arcadia, and the closely related Cypriot, spoken in the farther-flung island of Cyprus, and both closely akin to Mycenaean) and **Pamphylian** (from the southern Turkish coast; not a pure East Greek dialect, but a curious blend of Arcado-Cypriot, Aeolic, and Doric). The other group is **West Greek**, which contains dialects spoken generally to the south and west of East Greek: in the Peloponnese (Sparta), on many islands of the southern Aegean Sea, and in most of the Greek colonies in Sicily and Italy (where some of the oldest alphabetic Greek inscriptions hail from). Most of



Map 12.1 Greek dialects

the West Greek dialects belong to the group called Doric. Doric was traditionally the language of lyric poetry (aside from Sappho and the others mentioned above) and of the choral lyric passages in drama. One of the most important lyric poets for IE studies is Pindar, who reworked and incorporated much ancient material in his compositions. All the non-literary dialects are attested in alphabetic inscriptions except Mycenaean and Cypriot; the latter is written in a syllabary derived from Linear A.

Although Attic became the literary norm and is the most familiar Greek dialect, the other dialects are often more important for Indo-Europeanists, for they preserve various features that were altered or lost in Attic. We will note examples of this where relevant in the pages to follow.

The Greek alphabet

12.7. The early first millennium BC saw the rise of the Phoenicians, a people who originated in Lebanon but spread westward as highly successful maritime merchants. Their language belongs to the Semitic family and was closely related to Hebrew and Canaanite. They established outposts throughout the Mediterranean basin, the most famous of which was Carthage, and everywhere they went they brought not only goods for sale, but a far more precious item too – their alphabet. The Greek alphabet is an adaptation of the Phoenician alphabet, a historical fact that made its way into Greek folklore in the form of the alphabet's legendary Phoenician inventor, Cadmus (< Phoenician **qadm* 'east'). There are several versions of the early Greek alphabet, and where and when the Greeks first learned it from the Phoenicians is controversial. The oldest known examples of alphabetic Greek have been found in Euboea and Italy, and date from the early eighth century BC.

The regional Greek alphabets are mostly classifiable into an eastern and a western group. While a western alphabet used by Greek colonists in Italy is what gave rise to the Latin alphabet (see further §13.1–2), it was a form of the eastern alphabet that won out in Greece itself, slowly displacing the others beginning in the fifth century BC.

12.8. Like the related alphabets of modern Hebrew and Arabic, the Phoenician alphabet only had letters for consonants; the crucial innovation on the part of the Greeks was to use certain letters solely for vowel sounds. The Greek vowel letters A E I O Y (upsilon) and (in the eastern alphabet) H (eta) are adaptations of Phoenician letters for consonant sounds that were not present in Greek. One inherited letter, digamma (F), represented *w* and is not part of the canonical Greek alphabet, since *w* was lost early from the eventually standard dialect of Attic-Ionic; but digamma is common in inscriptions in other dialects. The last four letters of the Greek alphabet, Φ X Ψ Ω (*phi*, *khi*, *psi*, *omega*, standing for *ph*, *kh*, *ps*, *ō*), are additions.

From PIE to Greek

Greek and the other Indo-European languages

12.9. As mentioned in §10.4, many have thought that Greek, Armenian, Indo-Iranian, and the poorly known Phrygian form a subgroup within Indo-European

because of a number of mostly morphological features shared among them. Most of these features are shared between Greek and Indo-Iranian, but that may simply reflect the fact that these two branches are attested many centuries before the oldest Phrygian, and over a millennium before the oldest Armenian.

12.10. We may list some of the language's most noteworthy features differentiating it from the other branches. Greek is characterized by a triple reflex of the vocalized laryngeals, the loss of most word-final consonants, the devoicing of the voiced aspirates, and special developments of **i*. In morphology, the middle and the passive voice were differentiated in the future and the aorist, an aorist passive morpheme *-(th)ē-* was created, and the locative plural ending **-su* was replaced with *-si*.

Phonology

GREEK PHONOLOGICAL INVENTORY

12.11. Attic Greek of the fifth century BC had the following consonant inventory. There were four series of stops: the voiceless stops *p t k*, the voiced stops *b d g*, the voiceless aspirated stops *p^h t^h k^h* (usually transliterated *ph th kh*), and the voiced nasal stops *m n* and *ŋ* (the last only before a velar; pronounced like the *ng* of *sing* and written in Greek with a gamma, γ). There was also a voiceless sibilant *s*, a voiced affricate *dz* (zeta, ζ, also pronounced as a consonant cluster *zd*), and the voiceless glottal fricative *h*, which only occurred word-initially. There were two liquids, *l* and *r*, the latter always aspirated or voiceless in word-initial position; this was represented by the Romans as *rh-*, whence the spelling of such Greek-derived English words as *rhetoric* and *rhapsody*.

12.12. At the beginning of the Classical period, Attic Greek had five short vowels, phonetically [i e a o u], and seven long vowels, [iː eː aː oː uː εː οː], in addition to many diphthongs mentioned below. The two Greek long-vowel letters, Η and Ω, stand for the long vowels [εː οː]. Long [εː οː] were written *ei* and *ou*, the so-called "spurious diphthongs"; these digraphs (combinations of two letters representing one sound) were also used to represent what had originally been true diphthongs [ei] and [ou] from PIE **ei* and **ou*. Greek still had plenty of other diphthongs to go around: *ai ai̯ ēi oi̯ oi̯ ui̯; au eu̯ ēu̯ ou̯* (also *āu* outside Attic-Ionic); these also eventually developed into monophthongs.

During or before the early Classical period, perhaps as early as the sixth century, the high back vowel [u], short and long, became fronted to [y], as in French *une* or German *Brüder*. After this change, the long [oː] mentioned above developed further into [uː].

The complex early Classical vowel inventory was strongly reduced in the medieval period; see §12.55ff. below.

Development of the PIE consonants into Greek

STOPS

12.13. Greek is a centum language (§3.8), merging the PIE palatals with the plain velars. It continues the PIE voiceless and plain voiced series of stops unchanged, but devoiced the voiced aspirates. After the changes to the labiovelars (§12.15 below),

the language had the stops *p t k b d g ph th kh*. Though we think of *ph th kh* nowadays as the fricatives [f θ x], they were not pronounced as such until the early Christian era. Greek is the only language outside of Indic to preserve this series as aspirated stops.

12.14. Like Indic but independently, Greek shows the effects of Grassmann's Law, or aspiration dissimilation (§10.30). In an original sequence of two aspirated stops separated by one or more segments (**Th* . . . *Th*), the first lost its aspiration. Thus a reduplicated present like **dhi-dheh₁-mi* (< **dhe-dheh₁-mi*; cf. §5.24) first became pre-Greek **thi-thē-mi* and then, by Grassmann's Law, *títhēmi* 'I place'.

12.15. Among the stops, only the labiovelars have a particularly complicated history, and their development was different from dialect to dialect. The Attic-Ionic outcomes, illustrated in the tabular overview below, are dentals before an *e* (as well as *i* in the case of **k^w*), and labials before vowels and before most consonants.

	before <i>e</i>	before <i>i</i>	before back V or C
* <i>k^w</i>	<i>te</i> 'and' < * <i>k^we</i>	<i>tís</i> 'who?' < * <i>k^wis</i>	<i>póteros</i> 'which (of two)' (< * <i>k^wo-tero</i> -)
			<i>pémp₁tos</i> 'fifth' (< * <i>penk^wtos</i>)
* <i>g^w</i>	<i>adelphós</i> 'brother' < * <i>sm-g^welbhos</i> 'of the same womb'	<i>bíos</i> 'life' < * <i>g^wih₃os</i>	<i>boús</i> 'cow' < * <i>g^wōus</i> <i>ballō</i> 'I throw' (< * <i>g^wl₁-jō</i>)
* <i>g^wh</i>	<i>theinō</i> 'I kill' < * <i>g^when-jō</i>	<i>óphis</i> 'snake' < * <i>og^whis</i>	<i>néphō</i> 'I go without drink' < * <i>ne-eg^wh-ō</i> <i>phrén</i> 'midriff' < * <i>g^whrēn</i>

12.16. Before the decipherment of Mycenaean, it had been assumed that the labiovelars were still intact in Proto-Greek (the common ancestor of the attested Greek dialects) even though none of the dialects then known actually preserved them as labiovelars. This assumption had been made because it was the easiest way to explain certain consonant alternations within the dialects (such as Attic present *téll-omai* 'I turn' ~ aorist *e-pl-ómēn* 'I turned') and certain sound correspondences between dialects (such as Attic *tís* 'who?' = Thessalian *kís*; Attic *téssares* 'four' = Lesbian *péssures*). Mycenaean proved this hypothesis correct: it still preserved the labiovelars intact, or at least preserved them systematically as something different from labials, dentals, or plain velars. In transliterations of Mycenaean, they are represented with the letter *q*, as in *qe-to-ro-po-pi* (= *k^wetropopphi*) 'with four-footed (chattels)'.

12.17. All non-nasal stops were lost word-finally except in unstressed words like the preposition *ek* 'out of': **tod* 'the, this' (neuter) > *tó*, **ébbheret* 'he was carrying' (imperfect) > *éphere*, *hupódra* 'looking from underneath' < **upo-drk* (root **derk*- 'sec').

SIBILANT **s*

12.18. The sibilant **s* was generally preserved word-finally and next to stop consonants, as in *génos* 'kind' and *kákistos* 'worst'. Elsewhere it became *h*, as in *heptá* 'seven' < **septm*. This *h* had a tendency to disappear: it was lost between vowels except in Mycenaean (so pre-Greek **skelesa* 'legs' became *skeleha* in Mycenaean, written *ke-re-a₂*, with *a₂* representing *ha*) and was lost word-initially before

vowels in Ionic and some other dialects (a phenomenon called *psilosis*). The voiced variant *z in PIE that occurred before a voiced stop (§3.12) stayed intact in the sequence -zd-, which is written in Greek with the single letter zeta (ζ, z), as in *hízō* (phonetically *hízdō*) 'I set' < *síz-d- < reduplicated *sī-sd-.

12.19. Initial clusters beginning with *s- plus resonant lost the sibilant (again via an intermediate stage with h-), e.g. *niph-* 'snow' < *snig^wh-, cp. Eng. *snow*, Russ. *sneg*. This was a late development, as shown by metrical evidence in Homer: for example, the formulaic phrase *órea niphóenta* 'snowy mountains' scans as though it were still **órea sniphóenta*, which is what it must have been at the time of the incorporation of the phrase into the Homeric poems.

LIQUIDS AND NASALS

12.20. The liquids and nasals are preserved intact in Greek, except that final *-m became -n, as in the imperfect 1st sing. *épheron* 'I was carrying' < *ébherom. The nasal *n assimilated to the place of articulation of a following stop, as in *pémptos* 'fifth' < *penptos < *penk^wtos and *ánkura* 'anchor', pronounced [ank-] and written with gamma (ἄγκυρα).

12.21. By the historical period the syllabic nasals had lost their nasality and show up in most dialects simply as a: *pod-m 'foot' (accus. sing.) > pód-a; *n-dhg^whitom 'imperishable' > á-phthiton. Likewise, the syllabic liquids developed an epenthetic vowel a, either before or after the liquid: *k^wrd- 'heart' > kardíā (also *kradíē* in Homer); *p^wlth^w-u- 'broad' > platús. The syllabic liquids were still present in Common Greek, since the outcomes differ from dialect to dialect; there is also metrical evidence from Homer that *r was still *r at the time of the incorporation of certain phrases into the epics. Syllabic *r was possibly still preserved in Mycenaean.

GLIDES

12.22. The glide *j at the beginning of a word became either h- (e.g. *hág-ios* 'sacred' from PIE *jaǵ-, cp. Vedic *yájate* 'worships') or z- (e.g. *zugón* 'yoke' < PIE *jugom, cp. Eng. *yoke*). The difference has been attributed to the presence or absence of a preceding laryngeal, the standard account being that *H₁- > h- but *j- > z-; but not all are agreed that this is the case. Word-internally, *j disappeared, but had various effects on preceding consonants, on which see §12.26 below.

The glide *u was lost in standard Attic-Ionic, but in numerous other dialects it is preserved as w (digamma), as in Arcadian and Thessalian *woikos* 'house' (Attic-Ionic *oikos*; cp. Latin *uicus* 'homestead'). Although no digamma is used in the written transmission of the Homeric poems, the scansion of many lines demands that a w be read where it is etymologically expected. For historical purposes, Indo-Europeanists often cite Greek forms that once contained a digamma, or that are attested dialectally with one, with a parenthetical w, e.g. (w)oída 'I know', é(w)idon 'I saw'.

LARYNGEALS

12.23. Although there is much controversy concerning the details, Greek is of paramount importance for reconstructing the PIE laryngeals – arguably more important than any other language in the family. In the standard account, Greek uniquely has

three distinct descendants (the “triple reflex”) of the three laryngeals when they were vocalized: vocalized $*h_1$ $*h_2$ $*h_3$ became *e a o*, as in the passive verbal adjectives *thetós* ‘placed’ ($*dh_1$ -tó-), *statós* ‘stood, made to stand’ ($*sth_2$ -tó-), and *dotós* ‘given’ ($*dh_3$ -tó-).

12.24. These three outcomes are also found in word-initial position before consonants (usually resonants), as in *anér* ‘man’ from $*h_2nér$, *érebos* ‘darkness’ < $*h_1reg^w$ -, and *óphelos* ‘advantage’ < $*h_3bhel$ -. Until the advent of the laryngeal theory, the source of these initial vowels was not understood; they were simply labeled “prothetic vowels” – vowels tacked on. Much more limited vocalic reflexes of such laryngeals are found in Armenian and Phrygian (see §§16.21, 20.5).

12.25. Additionally, sequences of syllabic resonant plus laryngeal (the “long syllabic resonants”, §3.15) also have a triple reflex, depending on the laryngeal – another feature unique among the IE languages: $*ǵnh_1$ -tós ‘born’ > (-) *gnētos*, $*tʰh_2$ -tós ‘endured’ > Doric (-) *tlātos*, $*ǵnh_3$ -tós ‘known’ > (-) *gnōtos*. Sometimes, the outcome is a disyllabic sequence *eRe*, *aRa*, *oRo*, as in *génēsis* ‘birth’ < $*ǵǵh_1$ -ti- and *kámatos* ‘toil’ < $*kǵh_2$ -to-; this was probably the outcome when the original sequence was accented, but the matter is disputed.

CONSONANT CLUSTERS

12.26. A few of the major sound changes affecting consonant clusters and somewhat obscuring the IE state of affairs may be briefly mentioned. Most complicated, but also probably most widespread, are the changes that happened in consonant clusters ending in the glide $*j$, examples of which can be drawn abundantly from present stems of verbs formed with the suffix $*-je/o-$ (see §§5.29–30) and feminine derivatives in $*-ja$ (< $*-ib_2$, see §6.71). Selected examples of the outcomes in the Attic dialect are given below:

pre-Greek	Greek	pre-Greek	Greek
$*klep-jō$	<i>kléptō</i> ‘I steal’	$*tjeg^w-o-$	<i>sébomai</i> ‘I worship’
$*kī-āmer-$	<i>tēmeron</i> ‘today’	$*medhjos$	<i>mésos</i> ‘middle’
$*phulak-jō$	<i>phuláttō</i> ‘I guard’	$*melit-ja$	<i>mélitta</i> ‘honey-bee’
$*dīēus$	<i>Zeús</i> ‘Zeus’	$*ononm̥-jō$	<i>onomáinō</i> ‘I name’
$*nig^w-jō$	<i>nízō</i> ‘I wash’	$*leum̥-ja$	<i>léaina</i> ‘lioness’

As a side-note concerning English **etymology**, the Ionic outcome corresponding to Attic *-tt-* was *-ss-*, and sometimes English has borrowed the Ionic form of a word rather than an Attic one, such as the name *Melissa* (‘honey-bee’, vs. Attic *mélitta* above). Sometimes English borrowed both, as in *glossary* (Ionic *glōssa* ‘tongue’) but *polyglot* (Attic *glōtta*).

12.27. In East Greek, including Attic, the change of $*t$ to *s* happened not just before $*j$, but also before $*i$, as in the $*ti$ -abstract nouns. These nouns all end in *-sis* in Greek, whence many English borrowings like *basis* and *prognosis* (Gk. *básis*, *prógnōsis* < PIE $*g^w m̥$ -ti-, $*(-)ǵnh_3$ -ti-). This change is also seen in the 3rd pl. primary ending $*-nti$, which became first $*-nsi$; then the *n* disappeared, compensatorily lengthening the preceding vowel if there was one. Thus the thematic 3rd pl. active ending in Attic is *-ousi*, as in *phér-ousi* ‘they carry’ < $*pher-onsi$ < $*pher-onti$ (in Doric, the form is still *phéronti*).

12.28. The PIE clusters of two dentals (§3.36) became *st* in Greek, as in (*w*)*istōr* 'judge' < **uidtōr* 'knower' (of right and wrong).

Development of the PIE vowels into Greek

12.29. Greek preserves the PIE vowels and diphthongs (including the long diphthongs) more faithfully than any other IE language. Because of this, ablaut is reflected particularly clearly, since the ablauting vowels *e* and *o* are kept distinct from one another (in contrast to Indo-Iranian), and *o* is distinct from *a* (in contrast to Germanic, Balto-Slavic, and most of Indo-Iranian and Anatolian).

The short vowels are preserved intact in such words as *tís* 'who?' (**kʷis*), *pénte* 'five' (**penkʷe*), *hágios* 'sacred' (**ǵag-*), *nóstos* 'homecoming' (**nos-*), and *gónu* 'knee' (**ǵonu*). The long vowels (whether from original long vowels or from sequences of short vowel plus laryngeal) are preserved in such words as *íos* 'poison' (**uis-o-* or **uiHs-o-*), *patēr* 'father' (**ph₂tér*), Doric *hādús* 'sweet' (**suddú-* < **sueh₂du-*), *ókús* 'swift' (**ókús*), and *ikhthūs* 'fish' (**ǵhdhūs* or **ǵhdhuHs*). The one characteristic point on which Attic-Ionic innovated is in the treatment of **ā*, which in most positions or all shifted there to *ē* (e.g. *mētēr* 'mother' vs. Doric *mātēr* < **mātēr*).

DIPHTHONGS

12.30. Both short and long diphthongs are faithfully preserved in Greek: **ei* **ai* **oi* in *leípō* 'I leave' (**leikʷ-*), *lai(w)ós* 'left' (**laiwos*), and *léloipa* 'I left' (**le-loikʷ-*); and **eu* **au* **ou* in *rheúma* 'stream' (**sreu-mn-*), *haiós* 'dry' (**saus-o-*), and *akouō* 'I hear' (**h₂kous-*). By the fifth century BC, the diphthongs *ei* and *ou* had been monophthongized to [e:] and [o:], long vowels that already existed in the language because they had arisen from certain vowel contractions and other changes. (The spellings *ei* and *ou* for [e:] and [o:] are therefore often called "spurious diphthongs," as we saw in §12.12.)

Long diphthongs are continued in such forms as the dative singular of *o*-stems like *anthrōp-ōi* 'for the man'. Starting around the third century BC, however, the *i* in the long diphthongs *ai* *ei* *oi* ceased to be pronounced, whence the orthographic convention in Greek of placing the *i* underneath the long-vowel letter (α η ω). This convention is not followed in all modern texts, and postdates Classical Greek.

THE ACCENT

12.31. Ancient Greek, like Vedic Sanskrit and some Balto-Slavic languages, had a mobile pitch-accent. (Modern Greek, though, has a stress-accent like English.) In most dialects the position of the accent was not usually predictable except in verbs. Judging by ancient descriptions, the accented syllable had higher pitch than unaccented syllables, and is ordinarily marked with an acute accent (´). Syllables containing long vowels and diphthongs could have the high pitch on either the first or the second half (technically called the *mora*) of the syllable; accent on the first *mora* is denoted by a circumflex accent (˘). Thus *ái* is different from *ai*, the latter representing *ái* (high pitch on the first *mora*). Since Byzantine times it has been customary to write an acute accent on a final syllable as a grave (̀) when another word follows: *theá* 'goddess' but *theà* *Thétis* 'the goddess Thetis'. But there is no evidence that the grave represents a different kind of accent from the acute.

12.32. A small handful of particles, pronouns, and short verb forms are clitic and do not receive an accent; however, if they occur in groups of two or more, then all but the last one receive an acute, as in the string *gar te me phēsi* becoming *gár té mé phēsi* in the Homeric text (§12.65) at the end of the chapter. Some of the rules for the accentuation of clitic chains are strongly reminiscent of the rules for verbal accentuation; it may be that verbs themselves were weakly stressed or unaccented in the prehistory of Greek. This would be paralleled by the weaker phonetic status of verbs in Vedic and elsewhere in IE (§5.59).

Morphology

THE VERB

12.33. The morphology of the Greek verb is very close to that of Sanskrit and of reconstructed PIE. Verbs in Greek do not belong to “conjugations” as in a language like Latin (see §§13.12ff.); rather, for the old core of verbs, tense-stems are derived directly from roots. Greek has innovated chiefly by expanding the number of forms of the verb: it has a true future tense, and it has created a new passive, distinct from the middle, in the future and aorist. The Greek verb had up to six so-called principal parts or stems: the present (e.g. *gráphō* ‘I write’), a future (*grápsō* ‘I will write’), an aorist active or middle (active *égrapsa* ‘I wrote’), a perfect active (*gégrapha* ‘I have written’), a perfect mediopassive (*gégrammai* ‘I have written for myself, I have been written’), and an aorist passive (*egráphēn* or *egráphthēn* ‘I was written’). All tenses having secondary endings (imperfect, aorist, pluperfect) were prefixed with the augment *e-* in the indicative mood, although in Homer (and occasionally in other poets) this is sometimes lacking. Most of the time augmentless forms have no discernible difference in meaning from augmented ones.

12.34. Most IE present classes are still found in Greek, although the number of athematic verbs has been much reduced, with some types such as nasal-infix presents barely found. Two notable innovations in the present stems are the combination of reduplication with the suffix *-skō* (< **-skē/o-*; e.g. *mi-mnē-skō* ‘I remind’) and the double nasal presents with both nasal infix and nasal suffix *-anō* (e.g. *li-m-p-anō* ‘I leave’, Greek root *leip-*).

12.35. The primary personal endings. Primary active and middle endings in thematic and athematic verbs are exemplified below by the present indicative of the verbs *phēmí* ‘I say’, *didomai* ‘I give (for myself), I am given’, *phérō* ‘I carry’, *phéromai* ‘I carry (for myself), I am carried’. Dual endings are not given.

	Athematic		Thematic	
	Active	Mediopassive	Active	Mediopassive
sing. 1	<i>phēmí</i> ¹	<i>didomai</i>	<i>phérō</i>	<i>phéromai</i>
	2 <i>phéis</i> ²	<i>didosai</i>	<i>phéreis</i>	<i>phérei</i> ³
	3 <i>phēsi</i> ¹	<i>didotai</i>	<i>pherei</i>	<i>phéretai</i>
pl. 1	<i>phamén</i>	<i>didómetha</i>	<i>phéromen</i>	<i>phérometha</i>
	2 <i>phaté</i>	<i>didosthe</i>	<i>phérete</i>	<i>phéresthe</i>
	3 <i>phāsi</i> ³	<i>didontai</i>	<i>phérousi</i> ⁴	<i>phérontai</i>

¹ Doric preserves the original vowel in *phāmí*, *phātí*. ² Remade from **phēs*. ³ Doric *phantí*.

⁴ Doric has *phéronti*. ⁵ Homeric has *phéreai*.

12.36. The primary active athematic endings are easily derivable from PIE (see §5.12ff.) except for the 2nd sing. *-s*. The expected ending *-si* occurs only in the dialectal form *essí* 'you are' (and even here has been restored by analogy); elsewhere, the secondary ending *-s* seems to have infiltrated the primary paradigm. In the thematic conjugation, both the 2nd and 3rd singular are of uncertain origin; we would expect 2nd sing. **-ei* (< **-ehi* < **-e-si*) and 3rd sing. **-eti* (Attic **-esi*).

12.37. The primary middle endings show much more innovation than the active endings. To cite one example, the 1st sing. *-mai* is a refashioning of earlier **-ai* under the influence of the active ending *-mi*; this *-ai* in turn is a combination of the PIE 1st sing. middle ending **-h₂e* (> **-a*) and primary active particle **-i* (§5.13), which in Greek and several other branches replaced the original primary middle marker **-r*. The 3rd person endings *-tai* and *-ntai* are a replacement of *-toi* and *-ntoi* (ultimately PIE **(n)to* plus **-i*), which were still preserved in Mycenaean and Arcado-Cypriot (e.g. Arcado-Cypriot *keitoi* 'lies').

12.38. The secondary personal endings. Below is a chart illustrating the secondary endings, from the imperfect tenses of the verbs given previously: *éphēn* 'I was saying', *edidómēn* 'I was giving (for myself), I was being given', etc.

	Athematic		Thematic	
	Active	Mediopassive	Active	Mediopassive
sing. 1	éphēn	edidómēn ²	épheron	epherómēn ²
2	éphēs	edídoso	épheres	ephérou ³
3	éphē ¹	edídoto	éphere	ephéreto
pl. 1	éphamen	edidómetha	ephéromen	epherómetha
2	éphate	edídosthe	ephérete	ephéresthe
3	éphasan	edídonto	épheron	ephéronto

¹ Doric *éphā*. ² Doric *edidómān*, *epherómān*. ³ Earlier **ephereo* < **ephereso*.

12.39. The secondary active endings are mostly the predicted outcomes from the corresponding endings in PIE, except for the 3rd pl. athematic ending *-san*, which has been taken over from the *s*-aorist (see below). The secondary mediopassive endings have been refashioned since PIE times, just like their primary counterparts.

12.40. The perfect. The perfect is fairly well preserved, although the old alternation between *o*-grade and *zero*-grade has mostly been leveled out in favor of the *o*-grade, except in some archaic forms found especially in Homer. In the verb (*w*)*oída* 'I know', historically a perfect without reduplication but thought of as a present in Greek, the IE ablaut relationships are still clear: singular (*w*)*oída* (*w*)*oístha* (*w*)*oíde*, plural (*w*)*ídmen* (*w*)*íste* (*w*)*ísāsi* (Doric *ísanti*).

Greek also has three different pluperfect formations, none of which has a clear claim to being very ancient. A future perfect, formed to the perfect mediopassive stem, is also an innovation within Greek.

12.41. The aorist. Greek preserves all the PIE aorist categories. The thematic aorist in Greek, as in PIE, has the same endings as the imperfect, e.g. sing. *élip-on* *-es* *-e*, pl. *élip-omen* *-ete* *-on*, the aorist of *leípō* 'I leave'. The root and *s*-aorist deserve more attention; sample paradigms are reproduced below (*éstēn* 'I stood', *edómēn* 'I gave (for myself)', *élusa* 'I released', *elúsāmēn* 'I dissolved').

	Root aorist		s-Aorist	
	Active ¹	Middle	Active	Middle
sing. 1	éstēn	edómēn ²	élūsa	elūsámēn ²
2	éstēs	édou	élūsas	elúsō ³
3	éstē	édoto	éluse	elúsato
pl. 1	éstēmen	edómetha	elūsamen	elūsámetha
2	éstēte	édosthe	elúsate	elúsasthe
3	éstēsan ¹	édonto	élusan	elúsanto

¹ Doric has *ā* instead of *ē* and 3rd pl. *éstan*. ² Doric has *edómān*, *elūsāmān*. ³ Earlier **elūsao* < **elūsaso*.

12.42. In the root aorist, the Doric 3rd plural is noteworthy in preserving the zero-grade of the root (Doric *éstan*, from **e-sth₂-nt*, with vocalized laryngeal). The s-aorist has undergone some remodeling since PIE times, as in PIE it was athematic. The middle endings in the aorist are the ordinary secondary middle endings already discussed.

12.43. The aorist passive was usually formed with a suffix *-thē-* (e.g. *egráphthēn* 'I was written'), but in several older verbs the suffix was just *-ē-*. Curiously it had active personal endings and the participle is formed with the historically active suffix *-nt-* (e.g. *graphthént-* 'having been written'), and some archaic aorist passives such as *edáēn* 'I learned' and *ephánēn* 'I appeared' are not even passive in meaning. It may well be a descendant of the PIE stative in **-eh₁-* (§5.34).

12.44. The future. Most Greek futures are sigmatic, formed with a thematic suffix *-se/o-*: *lúō* 'I release' (present), *lúsō* 'I will release' (future). The source of the Greek sigmatic future is controversial, but most likely continues one of the PIE desiderative formations in **-s-*, probably **-h₂se-* (§5.37). Greek has invented a future passive by adding this *-se/o-* to the (unaugmented) aorist passive stem: aorist passive *e-timēthē-n* 'I was honored', future passive *timēthē-somai* 'I will be honored'. This is a fairly recent innovation; it is barely attested in Homer.

12.45. The subjunctive and optative. Greek has inherited the PIE subjunctive in both athematic and thematic verb stems, where it typically has modal force and not future force. Athematic stems in Homer still fairly commonly form "short-vowel" subjunctives that reflect the PIE subjunctive vowel **-e/o-*, as in Homeric *stē-o-men* 'let us stand' (aorist athematic stem *stē-*). In thematic stems, the contraction of the thematic vowel with the subjunctive vowel led to a subjunctive in long *-ō-* and *-ē-*, as in *phér-ō-men* 'let us carry' < **pher-o-o-men*.

The optative has also been inherited intact. The ablaut **-ieh₁-*/**-ih₁-* is preserved in athematic forms like the present optative 1st sing. *eīēn* 'may I be' (**h₁s-ieh₁-m*), 1st pl. *eīmen* 'may we be' (**h₁s-ih₁-me-*), and the expected optative of thematic stems, **-o-ih₁-*, is continued by Greek *-oi-* (with the *o*-vocalism generalized to all persons), as in *phér-oi-men* 'may we carry'.

12.46. Imperative. Greek has 2nd and 3rd person imperatives in the present, aorist, and perfect, in all voices. While several of these are new creations, the endings are mostly familiar from PIE. The athematic 2nd sing. **-dhi* is continued as *-thi* in a few athematic forms (e.g. *í-thi* 'go!'), while the thematic is endingless (*phére* 'carry!'). The 2nd pl. active ends in *-te* everywhere, as expected (*phérete* 'carry!').

12.47. **Infinitives and participles.** Infinitives and participles are very numerous in Greek, being formed from all the verb-stems. Greek active infinitives contain a suffix **-en*, **-men*, or **-ai*, sometimes in combinations like *-(e)nai* or *-menai*. In Attic, **-en* (probably < **-hen* < **-sen*) is found in thematic verbs, as in *phérein* 'to carry' < **pher-e-en* (cp. Mycenaean *e-re-e* 'to row' = *erehen*), while **-nai* is found in athematic verbs (*didó-nai* 'to give'). These endings are mostly creations within Greek, as are the peculiarly abundant participles, of which ten are formally distinguished. Except for the perfect, the active participles have a stem in *-nt-* which is inherited from PIE (such as the present *gráphont-* 'writing' and aorist *grápsant-* 'having written'), and except for the aorist passive, the middle and passive participles have a stem in *-meno-*, likewise inherited (such as the present mediopassive *graphómenos* 'being written' and future middle *grapsómenos* 'about to write for oneself'). The inherited perfect active participial stem **-yos-* has been refashioned to **-yot-* (with a *-t-* of disputed origin) except in Mycenaean, which continues the older form as *-woh-* (e.g. *te-tu-ko-wo-a₂* = *tetukhwoha*, neuter pl. of the Mycenaean perfect participle of *teúkhō* 'I fashion, craft').

THE NOUN

12.48. In contrast to the verb, which has gained a number of forms since PIE times, the number of Greek noun forms and stem-classes has shrunk somewhat. The number of cases has been reduced to six: **nominative**, **vocative**, **accusative**, **dative**, **genitive**, and **locative**. Most Greek grammars consider the locative a vestigial case, but it is alive and well in place-names and some locational adverbs, such as *oíkoí* 'at home'; its other uses, as with prepositions indicating location, have been taken over by the dative. The Greek dative has also taken over the uses of the instrumental, and in most dialects the genitive has absorbed the functions of both the genitive and ablative (compare the situation in Balto-Slavic, §18.30). Mycenaean may have still had a separate ablative, and a mostly instrumental ending *-phi* (continuing PIE **-bhi-*) is still productive in Mycenaean and Homer, but is not productive thereafter. It seems to have been indifferent to number (so *íphi* 'with might' [sing.], *naúphi* 'with ships' [pl.]), which has some analogues in Vedic Sanskrit (§6.17).

12.49. The following are the paradigms for the nouns *lúkos* 'wolf', *hórā* 'season' and *sōtēr* 'savior'.

	o-stem		ā-stem		Consonant stem	
	sing.	pl.	sing.	pl.	sing.	pl.
nom.	<i>lúkos</i>	<i>lúkoi</i>	<i>hórā</i>	<i>hōrai</i>	<i>sōtēr</i>	<i>sōtēres</i>
voc.	<i>lúke</i>	<i>lúkoi</i>	<i>hórā</i>	<i>hōrai</i>	<i>sōter</i>	<i>sōtēres</i>
gen.	<i>lúkou</i>	<i>lúkōn</i>	<i>hórās</i>	<i>hōrōn</i> ¹	<i>sōteros</i>	<i>sōtērōn</i>
dat.	<i>lúkōi</i>	<i>lúkois</i>	<i>hōrai</i>	<i>hōrais</i>	<i>sōtēri</i>	<i>sōtērsi</i>
acc.	<i>lúkon</i>	<i>lúkous</i>	<i>hōrān</i>	<i>hōrās</i>	<i>sōtēra</i>	<i>sōtēras</i>

¹ Homeric *hōrōn*.

12.50. The *o*-stem singular endings directly continue the relevant PIE endings discussed in §§6.45ff. The nominative plural in *-oi* is taken over from the pronominal declension, as in several other branches (§6.53). The endings of the *ā*-stems are mostly unremarkable; the singular in Attic shows the usual Attic change of **ā* to *ē*.

The *o*-stem and *a*-stem accusative plurals *-ous* and *-ās* arose by sound change from **-ons* and **-āns*, endings that are famously preserved intact in Cretan (the Doric dialect of Crete). Consonant-stem inflection is more straightforward than in PIE, since the endings no longer ablaut (though the stems sometimes do, e.g. *kuōn* 'dog', stem *kun-*; *patēr* 'father', stem *patr-*). The dative ending *-i* is from the PIE locative, which ousted the original dative in *-ei*, which is still in Mycenaean (see the Mycenaean text sample in §12.66 for an example and further discussion).

12.51. Greek still preserves the dual. The nominative-accusative ending **-h_i* is preserved intact in athematic nouns as *-e* (e.g. *pód-e* 'both feet'); the thematic ending **-oh_i* is also faithfully preserved as *-ō* (e.g. *hípp-ō* 'two horses').

Syntax

12.52. Because of its rich inflectional system, sentences in Greek, as in most other older IE languages, have many possible word orders. In poetry such as Pindar's, this freedom is stretched to its limits, but even here some theoretically possible word orders are not found, showing that certain orders were probably not grammatical. In particular, the order of clitic particles was constrained, as in other languages like Vedic and Hittite. The number of these particles was large; they conveyed various shades of meaning and are often difficult or impossible to translate. These particles, together with certain unstressed pronouns and conjunctions, tend to be grouped together near the beginning of the sentence after the first fully stressed word by Wackernagel's Law (§§8.22ff.). The rules of clitic placement, however, are not always that easy to divine, as they differ somewhat from author to author and dialect to dialect. In their essentials, though, their behavior is clearly comparable with clitic placement in Indo-Iranian, Celtic, and elsewhere.

12.53. A neuter plural subject takes a singular, and not a plural, verb, as in the phrase *tà zōia* (pl.) *trékhei* (sing.) 'the animals run'. This is also the case in Old Avestan and Hittite (cp. §6.68) and is an inherited feature.

Greek after the Classical Period

The Hellenistic period and the Koine

12.54. Toward the middle of the first millennium BC the prestige of Greek language and culture was extending outside the confines of the local Greek city-states. Both had been carried to the shores of Italy by colonists starting in the eighth century BC, where they would soon exercise considerable influence over the Etruscans and other indigenous peoples. Of even greater importance for the fortunes of Greek was the adoption of Attic-Ionic by around the fifth century as the official language of the court in Macedon, a kingdom to the north of Greece. When Macedon's power expanded in the fourth century under Philip II and his son Alexander III (the Great), Greek language and culture spread over an enormous area of the ancient world and ushered in the Hellenistic period, whose beginning is traditionally dated to the year of Alexander's death in 323 BC. During this time, Greek was spoken as far south as Egypt and as far east as Bactria. A uniform and somewhat simplified variety of spoken Greek called Koine (short for *hē koinē diálektos*, 'the common language') established itself as the medium of communication. It was based mostly on Attic but

had admixtures from Ionic and other non-Attic elements. As it became the standard administrative language, almost all the other Greek dialects died out, the one exception being the Doric dialect Laconian (living on nowadays as Tsakonian).

The Koine was used until around the reign of Justinian I in the sixth century AD; it is the language of the Septuagint and the New Testament, and was used by some authors such as the historian Polybius and the philosopher Epictetus. But most writers after the first century AD reacted against it and returned to pure Attic, including such familiar figures as Plutarch, Galen, Euclid, and Ptolemy. Mention should also be made of Hesychius, an Alexandrian glossator who compiled a lexicon of rare words probably in the fifth century AD. His lexicon, while sometimes faulty, contains a prodigious number of dialectal forms and forms from other languages that are only poorly known, and is of great value for Greek dialectology and IE studies.

12.55. Characteristics of the Koine. Most important for the later history of Greek were the changes that the Koine underwent in its phonology. First, the numerous distinctions among the vowels were reduced: *ei*, *ē*, and *i* all became pronounced [i]; *ai* and *e* merged as [e]; and *oi* and *u* merged as [y] (pronounced *ü* as in German). Over time, distinctions in quantity were lost; and the pitch-accent was replaced with a stress-accent. The aspirated stops *ph th kh* became fricatives [f θ x], and the voiced stops *b d g* became voiced fricatives [v ð ɣ]. In the morphology, the dual number and the optative mood were lost and the distinction between perfect and aorist was given up, with verbs retaining only one of the two past-tense stems (usually the aorist).

Byzantine Greek

12.56. The purist reaction alluded to above resulted in the adoption of an artificial archaizing literary language based on Classical Attic. This continued to be used as an administrative and literary language in the Byzantine period (roughly from the reign of the emperor Justinian I in the sixth century AD to the fall of Constantinople in 1453). The gap between the written and spoken languages widened with time, and as the Byzantine empire disintegrated and educational standards declined, spoken usages infiltrated the written language in ever larger numbers. Spelling errors (among other evidence) indicate that the two high vowels [i] and [y] merged as [i]. Additionally the dative case was lost in nouns, and the infinitive began to be replaced with other constructions found also in other languages of the region, such as Bulgarian; see further §19.5.

Modern Greek

12.57. Aside from Tsakonian in the eastern Peloponnese and the Greek dialects of Turkey, the varieties of Modern Greek are all descended from the Koine. Following the establishment of the new Greek state after the yoke of Turkish rule was cast off in 1828, a new literary standard called the *Katharevusa* ('purified') was created that was purged of foreign elements and drew on Ancient Greek for both vocabulary and inflections. The variety of Greek spoken in the Peloponnese became the spoken standard, known as *Demotic*. In 1976, Demotic also replaced the *Katharevusa* as the written standard, although the two varieties had essentially merged by then anyway; the result of their convergence is often called Standard Modern Greek (*Koini Neoelliniki*).

The Philology of Homer and Its Pitfalls

12.58. Because of the importance of the Homeric poems for the historical study of both Greek and PIE, it is essential to have some knowledge of the difficulties peculiar to these texts. In particular, the epics contain forms that are artificial and consequently of little or no historical value for comparative linguistics. We will present a sampling below.

12.59. The Homeric poems in their present form represent the accumulated labor of many generations of bards from different parts of eastern Greece. The result was a mixture of forms from different dialects and from different chronological stages. Each poet drew on a repertory of inherited and memorized formulaic poetic language, but in composing the epics in performance would inject newer material of his own devising. Bards constantly adapted the poetic language, and to make the verses scan they would sometimes create forms that from a historical point of view are wrong.

As noted earlier, the dialect of the epics is Ionic, with admixture from Aeolic and Mycenaean. Alongside Ionic forms like *téssares* 'four' and *hēmeis* 'we' are found the Aeolic equivalents *písures* and *ámmeis*. Aeolic and Ionic forms sometimes coexist in the same line, and there are even words that are artificial hodge-podges of morphemes from both Aeolic and Ionic, such as the dative plural *stēthessin* 'in (one's) chest', which has the Aeolic dative plural ending *-essi* plus an *-n* that many inflections have in Attic-Ionic.

12.60. Lines of Greek poetry were structured according to particular sequences of light (v) and heavy (–) syllables. A light syllable was one ending in a short vowel; all others counted as heavy. Syllable division ignored word-breaks; thus *ptolīethron épersen* 'he sacked the city' is syllabified *pto.li.eth.ro.ne.per.sen*. Heavy and light syllables are organized into basic units called feet, of which two occur in Homer: the dactyl (– v v) and the spondee (– –). A line of Homeric poetry contains six feet, most of which can be either dactyls or spondees, whence the name of the meter, *dactylic hexameter*.

12.61. Discrepancies between how a particular Homeric line is written and how it scans are common. Often, this is due to modernized spelling, in which older forms that did scan were replaced with newer forms that do not. For example, the *o*-stem genitive singular ended in disyllabic *-oio* in its oldest Homeric form (< PIE **-osjo*) and monosyllabic *-ou* in later Attic-Ionic. But there was also an intermediate stage, namely disyllabic *-oo* (two short syllables), a form that is nowhere preserved in the written tradition but whose former presence in some passages is betrayed by a mismatch between the spelling and the meter. Thus *Odyssey* 10.60 contains the phrase *Aiólou klutà dómata* 'famed dwellings of Aeolus', where *Aiólou* cannot scan as written because it would form the sequence – v –, which is not allowed in the dactylic hexameter. But restoring the genitive ending *-oo* solves the problem: *Aiólou klutà* (syllabified again without regard to word-breaks: *ai.o.lo.ok.lu.ta*) yields two dactyls, – v v | – v v.

12.62. The earlier presence of digamma (*w*) explains some other peculiarities. Normally in Greek poetry, a short vowel at the end of a word is elided (not pronounced) before a word beginning with a vowel; thus the sequence *me ôka* in line 416 in §12.65 below is written (and read) *m' ôka*. There are many exceptions to this in Homer, but usually they are due to the earlier presence of a digamma that blocked elision. An example is the phrase *Hēphaistoio ánaktos* 'of Hephaistos the

king', where if we restore the digamma that originally began the word for 'king', the irregularity disappears: (*)*Hēphaistoio wánaktos*.

12.63. A number of artificialities of the Epic language are due to so-called poetic "licenses." Poetic licenses are conceived as the conscious bending of rules of phonology or grammar to make a word fit the meter. The bards presumably could not do this willy-nilly; they took advantage of pre-existing and naturally occurring variants and then extended the properties of those variants analogically. One famous such license in Homer is poetic lengthening, the lengthening typically of the initial syllable of a word that could not otherwise have fit the meter, e.g. *Āpóllōna* 'Apollo' (accus. sing., for *Apóllōna*).

12.64. Finally, the alphabet used in the earliest stages of the written transmission of the epics did not distinguish short from long *e* and *o*, and words were written together without a break. These facts led to numerous errors on the part of copyists, including the creation of ghost-forms. An example is the story of the two nearly identical Greek adjectives meaning 'horrible, chilling', *kruóeis* and *okruóeis*. Both are descendants of PIE **kreuh₂-* 'gore', but the *o-* of *okruóeis* has no obvious source. Philological examination of the Homeric passages containing the problematic form *okruóeis* reveals two interesting facts: first, the word preceding *okruóeis* is always an *o*-stem genitive singular ending in *-ou* (e.g. in the phrase *epidēmíou okruóentos* 'of chilling civil [war]'); and second, this syllable must scan light to fit the meter, whereas normally (since it is a diphthong) it should scan heavy. Now, as just discussed above, the ending *-ou* is of somewhat recent vintage in Homer; it was earlier *-oo*, which is the ending we would expect to have been current at the time of the composition of much of the epics. If we combine the expected ending with the unproblematic form of the adjective, we would get the phrase *epidēmíoo kruóentos*, which happens to scan perfectly. For this scenario to work, all we need is an explanation for how *epidēmíoo kruóentos* became corrupted to *epidēmíou okruóentos*. The explanation is straightforward: since the earliest Greek texts were written without word-divisions, this would have been transmitted as *epidēmíookruóentos*, and later copyists who only knew the newer genitive singular ending *-ou* misparsed the sequence ... *ookru* ..., thinking that the second *o* was part of the following word and that the first *o* was meant to be the ending *-ou*. We conclude that there was originally just one adjective for 'chilling', *kruóeis*; the *o-* of *okruóeis* has no etymological significance, and the word entered Greek through a copyist's error. (Some authorities believe that this sort of error could already have arisen during the preliterate, oral transmission of the text through singers' misunderstandings rather than copyists' errors.)

Homeric Greek text sample

12.65. *Iliad* 9.410ff. Achilles describes his double fate.

- 410 mētēr gár té mé phēsi theà Thētis argurópeza
dikhthadías kēras pherémen thanátoio télosde:
ei mén k' aūthi ménōn Trōōn pólin amphimákhōmai,
óleto mén moi nóstos, atár kléos áphthiton éstai;
ei dé ken oĩkad' híkōmi phílēn es patrída gaian,
415 óletó moi kléos esthlón, epì dērōn dé moi aiōn
éssetai, oudé ké m' ōka télos thanátoio kikheíe.

- 410 For my mother, the silver-footed goddess Thetis, says to me
that (I) have a double fate towards the end of death.
If, on the one hand, staying here I fight beside the city of the Trojans,
my homecoming is lost, but I will have imperishable fame;
on the other hand, if I go home to my beloved fatherland,
415 noble fame is lost for me, but my life will be long,
and the end of death would not come swiftly to me.

12.65a. Notes. 410. *mētēr*: 'mother'; Doric *mātēr*. *gár té mé*: string of clitics placed after the first stressed word in the sentence (cp. Anatolian, §9.13): *gar* = 'for' (explanatory particle); *te* = 'and' (PIE **k^we*, also in Lat. *-que*, Skt. *ca*), connecting this sentence to the preceding and not translated here; and *me* is the accus. enclitic pronoun 'me', the subject of the following clause of indirect statement, where the main verb is an infinitive (cp. English *I consider him to be dishonest*). The pronoun has moved out of the clause of indirect statement and into the main clause (clitic raising). *phēsi*: 'says', clitic verb form; Doric *phātí*, PIE **bheh₂-ti*, cognate with Lat. *fā-tur* 'speaks'. *theā*: 'goddess' (**dhh₂-s-ā*); not related to Latin *dea* 'goddess', but rather to Lat. *fēstus* 'holy, festal' < **dheh₂-s-* 'sacred' and *fānum* 'temple' < **dhh₂-s-no-*.

411–12. *dikhthadías kēras*: 'double fates', object of *phérein* 'to carry, have', an Aeolic infinitive (Attic *phérein*). *thanátoio*: 'of death', with the archaic genit. *-oio* (§6.48). *télosde*: 'to the end, to completion'; *-de* indicates place to which (also line 414, *oíkad(e)* 'homewards'); it is probably related to Eng. *to*. *ei mén*: 'if, on the one hand . . .'; *k'*: elided form of *ke*, a particle (variant *ken* line 414) used in certain conditional clauses; maybe related to Hittite *-kan*, Ved. *kām* (§7.28). *ménōn*: 'remaining, staying', pres. partic. (stem *ménont-*), cognate with Lat. *maneō* 'I remain'. *amphimákhōmai*: 'I fight around, fight beside', pres. subjunc. mid.; *amphi-* comes from **h₂mbhi-* (whence German *um* 'around', Old High German *umbi*, and the *om-* of Eng. *ombudsman*).

413–14. *óleto*: 'is lost', 3rd sing. aorist middle of *ólumai* 'am lost'. *moi*: 'to/for me', clitic dat. 1st sing. pronoun; dative of possession ('my homecoming') or dative of reference ('as for me . . .'). *nóstos*: 'homecoming', a rare IE formation, with the **-to-* suffix added to the *o*-grade of the root (root **nes-* 'arrive home safely', also in Gk. *néomai* 'I come home safely' < **nesomai*, and the name *Nés-tōr*). *kléos áphthiton*: 'imperishable fame', directly from PIE **kleuos ṛdhg^hbitom* (see §2.37). This is its only occurrence in Homer. *éstai*: 'will be', shorter form of *éssetai* (line 416), future of the verb *es-* 'be', middle alongside the active present (cp. §5.6). *dé*: 'but', a particle functioning both as adversative and as sentence connector. *híkōmi*: 'I come, arrive at', pres. subjunc. In Attic the form would be *híkō*; the addition of the athenatic 1st sing. ending *-mi* is a Homeric feature. *philén es patrída gaian*: 'to (my) beloved fatherland'. The adjective (*philén*) has been moved out of the prepositional phrase, a common stylistic feature in the early IE languages (cp. Latin *magnā cum laude* 'with great (*magnā*) praise'; recall §8.7).

415–16. *epi dērōn*: 'for a long time'; *dērōn* is from earlier **duāron*. The second syllable of *epi* scans long, which can only be explained if the cluster *du-* was still present at the time of the composition of this line or phrase (the syllabification being *e.pid.ūē.ron*). The adjective comes from **duh₂-ro-*, whose zero-grade **dub₂-ro-* is seen in Lat. *dūrus* 'enduring'. *aiōn*: 'life, lifetime', borrowed into Eng. as (*a*)*eon*; ultimately < PIE **h₂oijn* 'life, life-force'; see also next. *oudé*: 'and . . . not, nor', combination of *dé* (line 414) and the negative *ou* 'not', also ultimately < **h₂oijn* 'life' (see §7.24). *ōka*: 'swiftly', an old adverbial accusative of the adjective *okús*, PIE **ōkú-* (Ved. *āsú-* 'swift', Lat. *ōc-ior* 'more swiftly').

Mycenaean text sample

12.66. One of the tablets found at Pylos, PY Ta 722, an inventory of vessels and furniture. "FOOTSTOOL" refers to a pictogram representing a footstool.

ta-ra-nu a-ja-me-no e-re-pa-te-jo a-to-ro-qo i-qo-qe po-ru-po-de-qe po-ni-ke-qe
 FOOTSTOOL I
 ta-ra-nu a-ja-me-no e-re-pa-te-jo ka-ra-a-pi re-wo-te-jo so-we-no-qe
 FOOTSTOOL I
 ta-ra-nu a-ja-me-no e-re-pa-te-ja-pi ka-ru-pi
 FOOTSTOOL I

One footstool inlaid with (a picture of) a man, horse, octopus, and palm tree (?) in ivory. One footstool inlaid with ivory lions' heads and grooves (?). One footstool inlaid with ivory nuts (?).

12.66a. Notes. 1. ta-ra-nu: *thrānus* 'footstool', otherwise only in Homer (*thrēnus*). a-ja-me-no: 'inlaid', mediopassive participle in *-menos* to an otherwise unknown verb. e-re-pa-te-jo: *elephanteiōi* 'made of ivory' (adj.), dat. sing. In the next line, *e-re-pa-te-jo* spells the dat. pl. *elephanteiois*, a-to-ro-qo: *anthrōqōi* 'man', dat. sing. The Myc. form shows that the *-p-* in the Classical form *ánthrōpos* comes from a labiovelar, but the source of this word is still disputed. i-qo-qe: *hiqqōi-q'e* 'and horse', dat. sing.; Classical *hippōi te*. The Myc. spelling system is like Latin in showing that this conjunction (= Lat. *-que*) was enclitic and pronounced together with the preceding form as a single word. The *h* and the *i* of Gk. *hippos* are irregular and still unexplained: we would expect **ep(p)os* (PIE **ekyos*). po-ru-po-de-qe: *polupodei-q'e* 'and octopus', dat. sing.; this and the following form are important in showing the inherited dat. sing. ending *-ei* (spelled *-e*), replaced in the rest of Greek (and in some Myc. texts) by the locative *-i*, except for trace survivals like the Cypriot name *Diwei-philos* 'dear to Zeus'. The word for 'octopus' is Classical *polúpous* (*polúpod-*), lit. 'many-footed'. po-ni-ke-qe: *phoinikei-q'e*, either 'and palm-tree' or 'and griffin', Classical *phoiniks*. There follows a pictogram of a footstool.

2. ka-ra-a-pi: *karābaphi* 'with heads', instr. pl., with Myc. instr. pl. ending *-phi*, surviving in Homer as an instr. ending in both sing. and pl. The first *a* may or may not be real; compare the two Homeric stem forms *karēat-* and *krāat-*. The forms are ultimately from PIE **kṛh₂-es* 'head' (> Ved. *śiras*). re-wo-te-jo: *lewonteiois* '(in the shape of) lions' (adj.), dat. pl. modifying *elephanteiois*. so-we-no-qe: 'and grooves'; probably the same word as Attic *sōlén* 'pipe, groove' < **sōh₂wēn*.

3. e-re-pa-te-ja-pi: *elephanteiāphi*, fem. instr. pl. ka-ru-pi: disputed; it could be an instr. pl. of an otherwise unknown *u*-stem noun **karu-* 'nut' (cp. Classical *káruon* 'nut'), or of an otherwise unknown **kalu-* 'flower-bud' (cp. Classical *káluks*). The inscription closes with a fourth line identical to the third.

Cretan text sample

12.67. Excerpt (V 29–45) from the Gortynian law-code, fifth or sixth century BC. Cretan is a Doric dialect with several interesting archaic features.

Ai de k' oi epiballontes oi men leiōnti datēththai ta krēmata, oi de mē, dikaksai ton dikastan epi toil leiōnsi datēththai ēmen ta krēmata tauta, prin ka dattontai. Ai de ka dikaksantos tō dikasta kartei enseiēi ē agēi ē perēi, deka statērāns katastasei kai tō krēios diplēi. Tnatōn de kai karpō kai wēmas kanpidēmas kēpipolaiōn krēmātōn, ai ka mē leiōnti datē[ththai - - ton dikast]an omnunta krinai porti ta mōliomena.

If some of the heirs at law want to divide the property but others do not, the judge shall decree that this property belongs to those wanting to divide it, until they divide it. And if, after the judge has made a judgment, anyone should rush in by force and

either drive or carry off (anything), he shall pay ten staters and double the value. And concerning livestock and produce and clothing and ornaments and movable wealth, if they do not want to divide (it) . . . the judge swearing by oath shall decide with reference to the pleadings.

12.67a. Notes (selective). First sentence. *oi*: 'the', masc. nomin. pl., Attic *hoi*. Cretan has lost *h-*. *epiballontes*: 'heirs at law', literally 'the ones devolving', a telescoping of a longer expression like 'the ones to whom the inheritance devolves'. *leiónti*: 'want', 3rd pl. pres. subjunctive of a Doric verb *leío* or *lō* 'I want', of unclear origin. Note retention of *-ti* in the ending (vs. Attic *-si* with loss of nasal, §12.27). *toil*: 'the', dat. pl., assimilated from *tois* to the initial *l-* of the following word (a Cretan, as well as a Homeric, feature). *leionsti*: 'ones wanting', dat. pl. pres. participle; note the retained cluster *-ns-*, characteristic of Cretan and some other Doric dialects. The dative is used to indicate possession. *datēththai*: 'to divide', pres. infinitive middle, with *-thth-* corresponding to *-sth-* in Attic-Ionic. *dattontai*: 'they divide', aor. subjunctive middle, with *-tt-* corresponding to Attic-Ionic *-s(s)-*.

Second sentence. *dikaksantos tō dikasta*: 'with the judge having decided', genitive absolute construction, corresponding to the ablative absolute in Latin. *kartei*: 'with force', Attic *krátei*. *statērans*: 'staters', a unit of currency, accus. pl., showing Cretan retention of *-ns*.

Third sentence. *tnatōn*: 'livestock', literally 'things that can perish', genit. pl.; cp. Attic *thnētōs* 'mortal', and the phrase *zōia panta thnētā kai phutā* 'all living things that die and grow', i.e. 'all living things, animals and plants', in Plato. *wēmas*: 'clothing', genit. sing., a neuter *n*-stem in the nomin.-accus. (*wēma* = Attic *hēma*, from **hes-mē*), but the genit. sing. is from an *ā*-stem **wēmā*. Compare Attic *gnōma* 'mark, token' alongside *gnōmē* 'opinion' (an *ā*-stem). The same is true of *anpidēmas* below. Note the preservation of initial *w-* in Cretan. *kanpidēmas kēpipolaion krēmatōn*: the first two *k*'s are from elided *kai* 'and'; *anpidēmas* (genit. sing.) would be Attic **amphidēmēs*, literally 'a thing that binds around', and *epipolaion krēmatōn* is the Cretan equivalent (genit. pl.) of Attic *ēpipla* 'utensils, movable property'.

omnunta: 'swearing', from *ómniimi*, cognate with Ved. *ámīti* 'swears' < **h₂emh₃-*. *krinai*: 'to decide', aor. infinitive, Attic verb *krínein*; root **krei-* 'to sift' as in Lat. *cribrum* 'sieve'. *portii*: 'towards, against, in reference to', Homeric *protí*, Attic *prós* (< **proti*, the sandhi variant before a vowel); cp. Ved. *práti*. *mōliomena*: 'pleadings', pres. mediopassive participle of *mōliō* (Attic *mōléō*) 'I plead'.

For Further Reading

The literature on the history of Greek is immense, and familiarity with the Greek alphabet is necessary for using any of it. The most exhaustive comparative grammar is Schwyzler 1939–71; more up-to-date, and the best recent historical grammar, is Rix 1992. In English one may consult Sihler 1995, the most recent combined Greek and Latin historical grammar; an earlier antecedent is the classic Meillet and Vendryes 1968. Two complete etymological dictionaries are available, one in French (Chantraine 1968–80) and one in German (Frisk 1960–72). The best overall description of the historical phonology of Greek is Lejeune 1972. A famous study of the pronunciation of Ancient Greek is Allen 1987, to which can now be added Devine and Stephens 1994, a fascinating if difficult work that attempts to reconstruct all levels of the prosodic phonology of Ancient Greek. Peters 1980 is a famous technical study of the history of the laryngeals in Greek. Useful for Indo-Europeanists is Risch 1974, a catalogue of Homeric forms arranged by formation and with historical commentary. Chantraine 1953–8 is indispensable for serious students of Homeric language, as are two exhaustive commentaries on the *Iliad* and *Odyssey* in English (Kirk et al. 1985–93 and Heubeck et al. 1988–92); the introductory essay to vol. 4 of Kirk 1985–93, by Richard

Janko, contains a very useful overview of the study of Homeric diction, poetics, and textual transmission. For the dialects, Buck 1968 is still eminently useful, though it lacks Mycenaean and some of the interpretations have been superseded. A more recently edited selection can be found in Duhoux 1984. Schmitt 1977 contains short but useful summaries of the dialectal developments, including Mycenaean. The standard compendium on Mycenaean, written by the decipherers themselves and including a history of the decipherment, grammar, edition of the entire corpus then known, and vocabulary, is Ventris and Chadwick 1973.

For Review

Know the meaning and significance of the following:

Mycenaean	Epic	psilosis
Linear B	Phoenicians	triple reflex of laryngeals
"Greek Dark Ages"	Grassmann's Law	Koine

Exercises

- 1 Discuss the fates of the following PIE sounds in Greek:

a * μ b * h_2 c * k^w d * s e * dh f * η

- 2 Using your knowledge of Greek sound laws, produce the Greek outcomes of the following PIE forms. Do not worry about accents or the Attic-Ionic change of * \tilde{a} to * \tilde{e} .

a * h_1le-h_1ludh- 'come' (perfect stem)	g * h_2ueh_1ti 'blows'
b * deh_3-rom 'gift'	h * $h_2i-h_2eus-oh_2$ 'I stay the night'
c * $h_2uoidenh_2$ 'song'	i * h_2enh_1mos 'wind'
d * $g^weibhus$ 'womb'	j * $bheidhoh_2$ 'I trust'
e * $tritas$ 'stretched'	k * $newos$ 'new'
f * $bhng^hus$ 'thick'	l * ng^en 'gland'

- 3 Demonstrate that Grassmann's Law in Greek was a separate change from Grassmann's Law in Sanskrit.
- 4 Using the following three forms, formulate rules for the outcome of the PIE labiovelars in front of nasals: $\acute{o}mma$ 'eye' < * $h_3ok^w-m\eta$, $amnós$ 'lamb' < ag^nos , $ophnís$ 'plowshare' < * uog^w-h-ni .
- 5 a The account given of Grassmann's Law in §12.14 is not quite complete. Expand it to account for the fact that pre-Greek * $si-sgh-\delta$ became Greek $ískhō$ 'I hold' and pre-Greek * $song^hā$ became Greek $omphē$ 'voice'.
- b Greek has a verb $ékhō$ 'I have', future $héksō$. Given your answer to 5a, and given the fact that the future of $gráphō$ 'I write' is $grápsō$, reconstruct what the present and future of 'have' would have been before the advent of Greek sound changes.

- c From the point of view of your expanded formulation of Grassmann's Law, what is curious about the word *theós* 'god' < **dhh₁s-os*?
- 6 a Given that PIE **syeh₂du-* 'sweet' became Gk. *hēdús*, PIE **sye* 'himself, herself' became Gk. *he*, and PIE **syekuros* 'father-in-law' became *hekurós*, what was the outcome of PIE **sy-* in Greek?
- b The second syllable of the phrase *phíle hekuré* 'dear father-in-law' (*Ilad* 3.172) scans long, which is unexpected for a short vowel in an open syllable. Provide a historical explanation for this scansion.
- 7 a Mention was made in §12.30 of the so-called "spurious diphthongs." Below are a number of examples of words containing the spurious diphthongs *ei* and *ou*. From this list, determine the sources of these diphthongs. The reconstructions are pre-Greek but not necessarily PIE. Ignore accents.
- | | |
|---|--|
| * <i>esmi</i> 'I am' > <i>elmf</i> | * <i>koruā</i> 'girl' > <i>kourē</i> (Epic, Ionic) |
| * <i>sems</i> 'one' (masc.) > <i>heīs</i> | * <i>enems₁</i> 'I doled out' > <i>éneima</i> |
| * <i>tons</i> 'the' (masc. accus. pl.) | * <i>eders₁</i> 'I flayed' > <i>édeira</i> |
| > <i>toús</i> | * <i>ksenyo_s</i> > <i>kseīnos</i> 'guest' (Epic, Ionic) |
| * <i>ktenjō</i> 'I kill' > <i>kteinō</i> | * <i>ghesr-ns</i> 'hands' (accus. pl.) > <i>kheīras</i> |
| * <i>phtherjō</i> 'I destroy' > <i>phtheirō</i> | * <i>oryos</i> 'boundary' > <i>oūros</i> (Epic, Ionic) |
| * <i>orsā</i> 'tail' > <i>ourā</i> | |
- b The verbal root **dyei-* 'fear' formed a perfect. What would the singular and plural perfect stems have been in PIE?
- c In Homer, two perfect forms from this verb are 1st sing. *deídoi_{ka}* (the *-k-* is a late addition) and 1st pl. *deídimen* (*-men* is the inflectional ending). Account for the first syllable *dei-*.
- d Account for the vocalism of the first member of the compound *theou-dés* 'fearing god' (contrast *theo-trephés* 'feeding the gods').
- 8 In Homer, the thematic verb *hépomai* 'I follow' conjugates in the singular and plural of the present as follows: (singular) *hépomai hépeai hépetai*; (plural) *hépómetha hépesthe hépontai*. The root in PIE and pre-Greek was **sek^w-*.
- a According to Greek sound changes, what ought the 2nd and 3rd singular and the 2nd plural have been?
- b Provide an explanation for the actually occurring forms of the 2nd and 3rd singular and the 2nd plural.
- 9 Below are several words or phrases from the Homeric epics and their metrical scansions, which are unexpected in one or more ways. For each phrase, indicate what is unexpected about the scansion and provide a historical explanation for it.
- a *órea niphóenta* 'snowy mountains' (*Odyssey* 19.338; *órea* = 'mountains', *niphóenta* = 'snowy') scans ~ ~ ~ ~ ~
- b *éban oīkónde* 'they went home' (*Odyssey* 1.424; *éban* = 'they went') scans ~ ~ ~ ~ ~

- c *es díphron árnas* 'to the chariot lambs' (*Iliad* 3.310) scans - - - - - (cp. lack of elision in *te amón* 'and of the lambs' at 4.158)
- d *parékhēi* 'provides' (*Odyssey* 19.113) scans - ∪ ∪ (don't worry about the last syllable; *par-* is a preverb meaning 'by'; use your answer to 5b to help you)
- 10 Provide a historical explanation of the form *-peza* '-footed' in *argurópeza* 'silver-footed', at the end of line 410 of the excerpt from Homer in §12.65. (Hint: see §6.71.)
- 11 Account for the boldfaced consonant clusters in the following words by positing a sound change not mentioned in the chapter.
- andrós* 'man' (genit. sing.; **h₂nr-ós*)
ámbrotos 'immortal' (earliest form **ḡ-mṛtos*, with dialectal change of **r* > *ro*)
- 12 The Greek thematic optative in the first person singular active ended in *-oimi* (e.g. *phér-oimi* 'may I carry'). The Arcadian dialect has a different ending, as in *ekselaun-oia* 'I would drive out'. Which of these endings, *-oimi* or *-oia*, is older from the IE point of view? Provide a historical explanation for both forms.

PIE Vocabulary IV: The Body

- **h₃ek^w*- 'EYE; to see': Lat. *oc-ulus*, OCS *oči* (dual)
 **h₃bhrū-* 'BROW': Ved. *bhrú-*, Gk. *ophrūs*
 **h₂ous-*, **h₂eus-* 'EAR': Gk. *oús*, Lat. *auris*
 **nas-* 'NOSE': Ved. *nas-*, Lat. *nāris*, OCS *nosŭ*
 **dn̥ghū-* 'TONGUE': Lat. *lingua* (Archaic *dīngua*), OIr. *teng*
 **h₁dont-* 'TOOTH': Ved. *dant-*, Gk. *odón* (*odont-*), Lat. *dēns* (*dent-*)
 **gonu-* 'KNEE': Ved. *jānu*, Gk. *gónu*, Lat. *genū*
 **pod-*, **ped-* 'FOOT': Ved. *pad-*, Gk. *poús* (*pod-*), Lat. *pēs* (*ped-*)
 **kerd-* 'HEART': Hitt. *kard-*, Gk. *kardiā*, Lat. *cor* (*cord-*)
 **esh₂r* 'blood': Hitt. *ēšhar*, Gk. *éar*, Latv. *asins*
 **dakru-* 'TEAR': Gk. *dákru*, Lat. *lacruma*
 **h₂ost-* 'bone': Hitt. *ḫastāi-*, Gk. *ostéon*, Lat. *os*
 **h₁org̃hi-* 'testicle': Av. *ərəzi-*, Arm. *orjī-*, Gk. *órkhis*, Alb. *herdhe*
 **ors-* 'ARSE': Hitt. *ārša-*, Gk. *ourá* 'tail'

13 Italic

Introduction

13.1. The Italic languages comprise most of the ancient Indo-European languages of Italy, as well as the modern-day descendants of one of those languages, namely Latin. Though Latin was destined to outshine its more obscure sister languages, its origins were humble. A casual observer of the linguistic and cultural milieu of early first-millennium-BC Italy would not have been led to predict greater fortunes for it than for its neighboring relatives; it was just one of a number of minor local languages in the center of the Italian peninsula spoken by predominantly pastoral tribes living in small agricultural settlements.

The Italic peoples were not indigenous to Italy, but arrived from the north probably by 1000 BC and slowly worked their way southward. North and central Italy had earlier been settled by successive waves of immigrants from across the Alps, while the southern regions, including Sicily, were partly under different cultural influence, being in contact with Aegean peoples to the east at least as early as the Sicilian Copper Age (c. 2500–2000 BC). Archaeological evidence points to widespread cultural exchange throughout the region, making it all the more difficult to link the known Italic peoples of historical times with specific prehistoric cultures.

By the early eighth century BC, Greek colonists from Chalcis in Euboea (northeastern Greece) had settled in Pithekoussai (modern Ischia, an island off the Italian coast by the Bay of Naples) and Cumae (on the coast near Naples), as well as other areas of the southern Italian coast and Sicily. These colonists brought a western Greek form of the alphabet with them (see §12.7), which soon spread rapidly into Italy. Alphabetic Greek inscriptions – among the oldest known – have been found on Italian soil dating to as early as 770 BC, and the first inscriptions in the non-Greek languages of Italy appear around the same time or even slightly before. (One recently discovered inscription from Gabii, a town near Rome, is already from c. 800 BC; it reads *euoi*, and although the language is uncertain, it is the oldest piece of alphabetic writing yet discovered that has vowels.)

13.2. Early on the Greeks came into contact with the southernmost settlements of the Etruscans, an ancient people whose homeland, Etruria (modern Tuscany), was located in northwestern Italy, bordering on Latium (to its south) and Umbria (to its east). Their culture was the dominant one in Italy from about the eighth century BC on; by the sixth century, Etruscan settlements flourished along most of the length of the western part of the peninsula. From the Greeks the Etruscans picked up the alphabet and many other aspects of Greek culture. The Etruscans never developed a

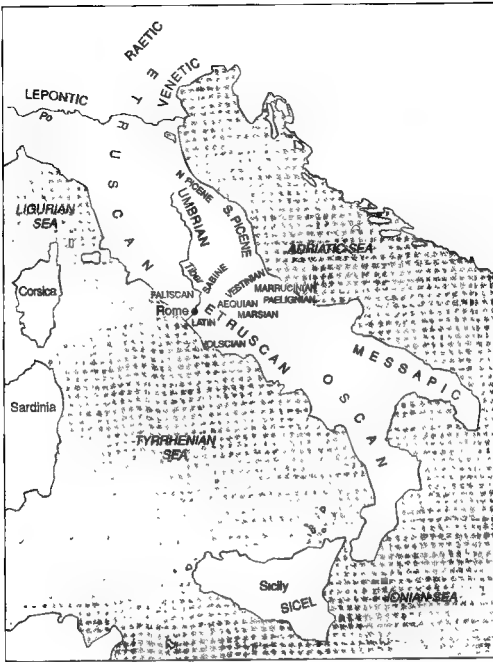
centralized state, but their impact on the fledgling civilization of the Romans was tremendous. Among other things, the Romans and many other peoples of Italy may have the Etruscans to thank for their knowledge of the alphabet. As Rome's power grew, the culture of the Etruscans waned, disappearing by about 100 BC; but Etruscan continued to be cultivated as a priestly language for several hundred years more.

The Etruscan language was not Italic, nor even Indo-European. Beyond this we know very little about it: although there is no dearth of inscriptions (over 13,000 have been found), they are mostly short dedicatory or funeral inscriptions containing only proper names, and the few longer texts are not well understood. The origin of the Etruscans, in fact, has been controversial since ancient times, with one school claiming they came to Italy from Asia Minor and another claiming they were a pre-Indo-European people indigenous to Italy. On the one hand, the testimony of the Greek historian Herodotus (who claimed they came from Anatolia), certain cultural practices (such as divination by inspection of the liver of a sacrificial animal, or hepatomancy), and the presence of some Anatolian loanwords in Etruscan support the theory of an origin from Asia Minor. On the other hand, there is no archaeological evidence of a migration, and Etruscan culture seems to grow organically from an earlier culture of north-central Italy called Villanovan. The existence of Lemnian, a language very similar to Etruscan and preserved on a stele and a few minor vase inscriptions on Lemnos (an island off the coast of Asia Minor) from the seventh or sixth century BC, does not settle the issue: Lemnian appears to be just a dialect of Etruscan, meaning it may have been simply the language of a group of Etruscan colonists from Italy. Aside from Lemnian, the only possible relative of Etruscan is Raetic, spoken in extreme northern Italy and neighboring areas.

13.3. Latin and the languages most closely related to it (of which only Faliscan is known) form the **Latino-Faliscan** branch of Italic and were spoken originally in a small region of west-central Italy south of Etruria. The languages comprising the other subbranch of Italic, **Sabellic** (also known as **Oscan-Umbrian**), were spoken over a considerably larger area of central and later southern Italy. Given the significant differences between the two linguistic groups already in the sixth century BC, Italic linguistic unity probably ended at the close of the second millennium BC. Some scholars have given these differences even greater weight and reject the notion of a single Proto-Italic language; according to this opinion, Latino-Faliscan and Sabellic belong to two separate branches of IE that partially converged due to later mutual influence. But this is a controversial view, and we shall assume a Common Italic stage in all that follows.

13.4. The seventh century BC saw the first inscriptions in Latin, and by the end of the sixth century the alphabet had spread eastward through Italy to the other coast, where the Piceni lived (§13.72), and northeastward into the territory of the Veneti (see below). The Samnites (whose language we call Oscan) and the Umbrians probably had the alphabet at this point also, but our earliest records in these languages come a bit later. As is customary in the field of Italic philology, when citing forms from a non-Latin inscription in a local non-Latin alphabet, boldface will be used. Inscriptional Archaic Latin forms will be given in small capitals.

In the extreme northeast of Italy lived a people known in ancient times as the Veneti, along the northern and northwestern shores of the Adriatic Sea. Their language, Venetic, bears affinities with the Italic languages but is not known well enough to be classified. It is treated in chapter 20.



Map 13.1 Languages of ancient Italy

"Italo-Celtic"

13.5. Italic shares several innovative features with Celtic, such as the o-stem genitive singular in **-i* (e.g., Lat. *uir-i* 'of a man', Ogam Irish *maq(q)i* 'of the son'), a superlative suffix **-mmo-* (as in Lat. *inf-imus* 'lowest', Celtiberian personal name *Uxama* 'highest'), the assimilation of the sequence **p...k^w* to **k^w...k^w* (as in the word for 'five', **penk^we*, which became Latin *quinque* and Old Irish *cóic* from **k^wenk^we*), and a subjunctive morpheme **-ā-* (as in Archaic Latin *fer-ā-t* 'he may carry', Old Irish *beraid* 'he may carry'). These shared features have led many Indo-Europeanists to posit an "Italo-Celtic" subgroup or dialect area of Indo-European. However, the hypothesis of an Italo-Celtic unity has never gained universal approval.

From PIE to Italic

Phonology

13.6. Stops. Italic is a *centum* branch, and therefore the palatal stops and the plain velars fell together as plain velars, as in the word *centum* itself, Latin for '100', from PIE **k₁ntom* (the *c* in the spelling of the word was pronounced *k* in pre-imperial Latin, still surviving today in Sardinian: see §13.49 below). The remaining PIE plain voiceless and voiced stops remained unchanged in Italic. The labiovelars have divergent developments in Latino-Faliscan and Sabellic (§§13.24 and 58).

The characteristic look of the Italic languages is due partly to the widespread presence of the voiceless fricative *f*, which developed from the voiced aspirates. Broadly speaking, *f* is the outcome of all the voiced aspirates except in Latin; there, *f* is just the word-initial outcome, and several other reflexes are found word-internally.

13.7. Dental-plus-dental sequences in the parent language became **ss* in Italic, as in Lat. *fissus* 'split' < **bhid-to-*. The consonant cluster **-tl-* became **-kl-*, as in Osc. *puklum* (accus. sing.) 'son' < **putlo-* and Lat. *pōculum* 'cup' (earlier *pōclum*, which is how the word scans in early Latin poetry) < **pōtlom*.

13.8. Resonants. The consonantal liquids, nasals, and glides remain unchanged: Osc. *anamúm* 'breath' (**h₂en_hmo-*), Lat. *alter* 'the other' (**al-tero-*), *iugum* 'yoke' (**iugom*), Umbr. *uiro* 'men' (**uir-o-*). Syllabic liquids developed a prothetic *o* before them (or *ol*), whereas syllabic nasals developed an *e* (*em en*): Lat. *cord-* 'heart' (**k₁rd-*), *mollis* 'soft' (**m₁ld-_{ui}-*), *septem* 'seven' (**septm₁-*), *nōmen* 'name' (**h₁neh₃m₁n₁-*). The outcome of syllabic nasals in word-initial syllables in Sabellic may have been different, however. According to a widely held view, the outcome there was *an* and *am*, as in Osc. *fangvam* (accus.) 'tongue' (cp. Lat. *lingua*, Archaic Lat. *dingua*, PIE **d₁ŋghu-*) and Umbr. *ander* 'between' (cp. Lat. *inter*, PIE **nter*). This view is not universally accepted, as the relevant forms have alternate explanations and it adds some unnecessary and unlikely complications to Italic historical phonology.

13.9. Sibilants. The sibilant **s* disappeared word-initially before a resonant, as in Lat. *nix* (stem *niu-*) 'snow' (vs. Eng. *snow* with preserved *s*-). Later in the history of Latin and Umbrian, *s* in certain positions was voiced to **z*, which then became *r* by a process known as *rhotacism* (see §§13.30, 13.65). Before consonants, this **z* disappeared with compensatory lengthening of the preceding vowel, as in the word for 'nest', Lat. *nīdus*, from PIE **nīzdos*.

13.10. Laryngeals. The laryngeals were lost in their non-vocalized form, but the vocalized laryngeals are preserved as *a*, as in Lat. *status* 'stood' (**sth₂-to-*), *datus* 'given' (**dh₃-to-*), and Osc. *anamúm* 'breath' (**h₂en_hmo-*; in Lat. *anima* 'breath' the internal vowel weakened to *-i-* as per §13.32). The outcome of syllabic resonant plus laryngeal was *Rā* or *aRa*, the latter probably when under the accent: Lat. *g₁nātus* 'born' (**g₁nh₁-tō-*), *strātus* 'strewn' (**st₁h₃-tō-*), and *palma* 'palm' from earlier **palama* (**p₁h₂-meh₂-*, cp. Gk. *palāmē*).

13.11. Vowels. The vowels and diphthongs were all preserved intact and are kept most faithfully in Oscan. In Latin, they subsequently underwent significant modification; see §§13.32ff. The mobile pitch-accent of PIE (§§3.30ff.) was replaced by a stress-accent on the first syllable of the word in Italic. This situation still obtained in early Latin, but was later replaced by the classical stress pattern (see §13.36 below).

Verbal morphology

13.12. Of considerable interest are the substantial innovations in the verbal system, which will be described here at some length, with certain details left for later. Broadly, Italic reorganized (1) the PIE present formations into a neat system of four conjugations, each characterized by a particular stem-vowel to which the personal endings were added; and (2) the PIE tense-aspect system into one based on the aspectual opposition between imperfective and perfective, with each of these having a future and a past tense.

THE FOUR CONJUGATIONS

13.13. With few exceptions, all Italic verbs belong to one of four classes or conjugations characterized by the stem-vowels *ā*, *ē*, *e*, and *i*. Compare:

Conj.	Latin	Oscan	Umbrian
I	<i>port-ā-re</i> 'to carry'	<i>faam-a-t</i> 'commands'	<i>port-a-tu</i> 'let him carry'
II	<i>hab-ē-re</i> 'to have'	<i>fat-i-um</i> 'to speak'	<i>hab-e</i> 'has'
III	<i>sist-e-re</i> 'to place'	<i>did-e-st</i> 'he will give'	<i>sest-e</i> 'you place'
IV	<i>aud-i-re</i> 'to hear'	<i>sakruv-i-t</i> 'consecrates'	<i>her-i</i> 'wants'

The rise of the conjugational system was due almost entirely to sound change. (All the examples in the following discussion are taken from Latin.) The first conjugation stem-vowel *-ā-* comes primarily from contraction of the sequence **-ā-je-*, which itself comes from various sources: denominative verbs from *ā*-stem nouns and adjectives (e.g. *cūrā-re* 'to take care' from *cūra* 'care'), factitive verbs from *o*-stems (e.g. *nouā-re* 'to make new' < **new-eh₂-je-*, cp. Hitt. *new-ah₂-*), and **-je/o*-presents from other types of verbs (e.g. *tonā-re* 'to thunder' < **(s)tonh₂-je-*).

The second conjugation stem-vowel *-ē-* comes from the contraction of the causative suffix **-éje-* (§5.32, e.g. *mon-ē-re* 'warn' < **mon-éje-*, literally 'cause to think') and from the stative suffix **-eh₁-je-* (§5.34, e.g. *alb-ē-re* 'to be white' and *sed-ē-re* 'to be sitting').

The third conjugation consists mostly of old thematic presents of various kinds (e.g. *ag-e-re* 'to drive, do' < **h₂eg₂-e-*; *bib-e-re* 'to drink' < **pi-ph₂-e-*; *albēsc-e-re* 'to grow white' < **albh-eh₁-ske-*), as well as a number of originally athematic presents that later became thematized, such as *find-e-re* 'split' < **bhi-n-d-*.

Presents formed with the suffix **-je/o-* added directly to the root have a more complicated history than the other kinds of **-je/o*-presents treated above. Such presents wound up in the fourth conjugation if the root they were added to was heavy, that is, ended in a consonant cluster, had a long vowel or diphthong, or consisted of two syllables (e.g. *sanc-i-re* 'to ratify' < **sank-je-*, *sepel-i-re* 'to bury' < **sepel-je-*). Otherwise they fell into a special class of the third conjugation that will be treated in §13.39 below, such as *fug-e-re* 'flee' < **bhug-je-*.

The rest of the fourth conjugation consists of denominatives from both *i*-stems and *o*-stems (the first e.g. in *moll-i-re* 'to soften' < *mollis* 'soft'; the second e.g. in *seru-i-re* 'be a servant (to), serve' < *servus* 'slave').

13.14. Very few athematic verbs still inflect athematically in Italic; relics include the Latin verbs *es-se* 'to be', *i-re* 'to go', and *uel-le* 'to want'.

THE ITALIC TENSE-ASPECT SYSTEM

13.15. The Italic tense-aspect system was based on an opposition between imperfective and perfective, each having a future and a past tense. The basic imperfective tense was the present (Lat. *portō* 'I carry, am carrying'), whose corresponding past tense was the imperfect (*portābam* 'I was carrying, used to carry') and whose future was the ordinary future (*portābō* 'I will carry'). The basic perfective tense was the perfect (*portāvī* 'I carried, have carried'), whose past tense was the pluperfect (*portaueram* 'I had carried') and whose future was the future perfect (*portāuerō* 'I shall have carried'). Each of these tenses except the two futures was also fitted out with a subjunctive.

THE IMPERFECT AND FUTURE

13.16. The PIE imperfect was lost without a trace in Italic. A suffixal morpheme **-f-*, an Italic invention derived from the PIE root **bhuH-* 'be, become', was used to form the imperfect and (in Latino-Faliscan only) the future. The imperfect used the stem **-fā-*, while the future used the stem **-fe-* (probably from an old subjunctive **bhu(H)-e-*; see directly below): Lat. imperfect *port-ā-bat* 'he was carrying', Umbr. *fu-fans* 'they were'; Lat. future *port-ā-bit* 'he will carry', Fal. *care-fo* 'I will do without'. The future is also formed (in Sabellic exclusively, and in Latin vestigially) with the suffix *-s-*, e.g. Archaic Lat. *fax-ō* 'I will do' (**fak-s-*), Osc. *dide-s-t* 'he will give', Umbr. *prupeha-s-t* 'he will purify before'. This is a continuation of one of the PIE future or desiderative formations in **-s-* (§§5.35ff.). Finally, the old PIE subjunctive, where it survives, became a future in Italic: for example, the PIE thematic verb **h₂eĝ-e-* 'drive' formed a subjunctive with stem **h₂eĝ-e-e-*, which became the future stem *ag-ē-* 'will drive' of the Latin verb *agere* 'to drive'. Note also the future *erit* 'he will be' (Archaic Lat. *esed*) from the athematic subjunctive **h₁es-e-* (cp. Ved. *ásati*).

The Italic imperfect in **-fā-*, as well as forms like the Latin imperfect stem *erā-* 'was', attest to the presence of a formation known as the *ā*-preterite. Its origins are unclear, though it has been compared to scattered past-tense forms elsewhere in the family that have a stem in *-ā-*, such as Doric Greek *erruā* and Old Church Slavonic aorists like *sūpa-* 'slept'.

THE PERFECT SYSTEM

13.17. The Italic perfect is a conglomeration of the IE perfect and aorist. All the Italic languages have a reduplicated perfect (Lat. *dedit* 'he gave', Fal. *peparai* 'I gave birth to', Osc. perfect subjunctive *fefacid* 'he might make', Umbr. *dede* 'he gave'), a non-reduplicated or de-reduplicated perfect (Osc. (kúm-)bened 'agreed', Umbr. *benust* 'will have come', Lat. *tulī* 'I brought'), and a long-vowel perfect (Lat. *ēgit* 'he drove, did', Osc. *hipid* 'had' [*< *hēb-*], Umbr. (pru-)sikurent 'they will have announced'). The old stative meaning of the perfect is still visible in such forms as Lat. *memini* 'I remember' (PIE **me-mon-*), but as a rule the Italic perfect is a past tense.

13.18. The pluperfect and future perfect were both formed with a morpheme **-s-*. The pluperfect, attested only in Latin, can be exemplified by Lat. *fu-eram* 'I had been' from earlier **fu-isam*. The future perfect, which happens to be particu-

larly well attested in Umbrian, can be exemplified by Lat. *fu-erit* (< earlier **fu-iset*) and Umbr. *fu-s-t*, both meaning 'he will have been'.

THE SUBJUNCTIVE

13.19. The Italic subjunctive is not a continuation of the PIE subjunctive, which became a future (see above). There are at least three subjunctive morphemes found in Italic, of which one continues the PIE athematic optative and the other two are of unknown origin. Their distribution is complex and need not be entered into here in detail; a few examples will suffice. The PIE optative morpheme is seen, with ablaut still intact, in Archaic Lat. *siēs* 'may you be' (**h₁s-ieh₁-s*), pl. *sītis* 'may you [pl.] be' (**h₂s-ih₂-te-*); but the zero-grade *-ī-* was generalized elsewhere. The so-called *ā*-subjunctive, also found in Celtic, is seen for example in the Lat. pres. subj. *habe-ā-s* '(that) you have' and the Osc. pres. subj. *pūtī-a-d* '(that) he be able'. Finally, the *ē*-subjunctive, which has no sure analogues outside Italic, is seen for example in the Lat. imperfect subj. *es-sē-s* 'you would be' and the Umbr. perfect subj. *herii-ei* 'he should want'.

PERSONAL ENDINGS

13.20. The old distinction between primary and secondary personal endings has traces in the third person, such as in Faliscan *fifiqod* 'they fashioned', with *-od* from secondary **-ont* and not primary **-onti*. The dual has disappeared. Like Anatolian, Tocharian, and Celtic, Italic generalized PIE **-r* as the marker of the mediopassive (called simply the passive in Italic linguistics). Some verbs, such as *sequor* 'I follow', inflected only in the passive, and are called "deponent" in traditional Latin grammar; these are often the descendants of PIE middle verbs (see §5.5).

PARTICIPLES AND INFINITIVES

13.21. The IE present active participle in **-nt-* is well preserved, as in Lat. *port-ant-* 'carrying', but only traces are found of the mediopassive participle in **-m(h₁)no-* (such as *alumnus* 'nursling, foster-son', literally 'one [being] nurtured'). The PIE infinitive in **-dhi-* (§5.54) is found in Sabellic (see below §13.61). The other Italic active infinitives, *-se* in Latin (e.g. *es-se* 'to be') and **-om* in Sabellic (e.g. Umbr. *er-om* 'to be'), descend from nominal formations and are not ancient. The Italic perfect passive participle in **-to-*, such as Lat. *cap-tus* 'taken' and Umbr. *uirseto* 'seen' (< **uid-ē-to-*), directly continues the PIE verbal adjective in **-tó-* (§5.57). Italic also created a future passive participle or gerundive in **-nd-*, of unclear origin: Lat. *dēlendus* '(about) to be destroyed', Osc. *ūpsannam* 'to be done'.

Nominal morphology

13.22. All the PIE nominal stem-classes are preserved in Italic, including even traces of the archaic *r/n*-stems (§6.31), as in Lat. *femur* 'thigh', stem *femin-*, and Umbrian *utur* 'water', stem *utn-*. The Latino-Faliscan and Sabellic branches differ in certain details of nominal inflection, but as an aggregate they show that Italic inherited the PIE case-endings with little change. The instrumental had been lost by the historical

period (though examples still survive in adverbial use, such as Lat. *bene* 'well') and its functions taken over by the ablative. The functions of the locative, too, were mostly taken over by the ablative, though it still survives as a productive separate case in place-names and certain nouns, e.g. Lat. *Rōmae* 'in Rome', *rūri* 'in the country', Osc. *meīai vīai* 'in the middle of the road'. The final dental of the o-stem ablative singular ending **-ōt* (§6.49) spread to the ablative of all the declensions: Osc. *toutad* 'by the people', Archaic Lat. *magistratud* 'with the office of a magistrate'. There is no dual.

13.23. As noted above, a peculiarity that Latino-Faliscan o-stem nouns share with Venetic, Messapic, and Celtic is a genitive singular ending **-i* (e.g. Lat. *uir-i* 'of a man', Fal. *Marci* 'of Marcus'), which is of uncertain origin. Alongside this, however, the more familiar genitive in **-osjo* was inherited as well, as in Archaic Latin *Popliosio Valesiosio* 'of Publius Valerius'. Neither occurs in Sabellic.

Latino-Faliscan

13.24. The Latino-Faliscan subbranch of Italic comprised Latin and Faliscan. In the mid-first millennium BC they were neighboring languages in a small area of west-central Italy. Because of the few remains in Faliscan, Latin is usually our only witness for Latino-Faliscan innovations. One difference between Latino-Faliscan and Sabellic that is immediately diagnostic is the divergent treatment of the labiovelars: they became labial stops in Sabellic (**k^w > p*, **g^w > b*) but not in Latino-Faliscan (**k^w* remained, and **g^w > ɣ*). Thus for instance **k^we* 'and' became Latin *-que* and Faliscan *-cue*, but Osc. *-pe*. See also §13.58.

Latin

13.25. Latin derives its name from Latium, a region of west-central Italy cut through by the lower part of the river Tiber as it flows westward to the Tyrrhenian Sea. One of the tribes in this area, around the Alban Hills, were the Latini, who eventually became dominant in central Italy and beyond.

The period of Latin from the earliest inscriptions to about the mid-second century BC is called Archaic Latin (also Old Latin). The first scattered inscriptional remains appear in the last quarter of the seventh century BC; inscriptions become relatively copious only in the third century BC, a century that also marks the beginning of preserved Latin literature. The earliest surviving literary fragments come from Livius Andronicus (born c. 284 BC); he is traditionally credited with being the first to set the Latin language to Greek meters. About two generations later came the comic playwright Plautus (254?–184? BC), the first author whose works survive in considerable quantity; his plays are followed by those of Terence (c. 195–159 BC). Also important for this time are such poets as Ennius, Accius, and Lucilius; of their works we unfortunately possess only single lines or short passages quoted by later writers. Latin prose begins with Cato the Elder (234–149 BC), whose book on agriculture, *Dē Agrī Cultūrā*, is of immense value for historians of Latin language, culture, and religion.

13.26. The Archaic period was followed by the period of Classical Latin, traditionally divided into the Golden and Silver Ages. The Golden Age, lasting until the

death of the poet Ovid in AD 17, saw for example the orations and other works of Cicero; the military commentaries of Caesar; the histories of Livy; and the poetry of Lucretius, Catullus, Horace, and Vergil. The Silver Age, dating until the death of the emperor Marcus Aurelius in 180, contains such literature as the tragedies and philosophical writings of Seneca; the novel *Satyricon* of Petronius; the *Natural History* of Pliny the Elder; the satires of Juvenal; and the histories of Tacitus and Suetonius. After this period came Late Latin, during which a large amount of early Christian literature was written, as by St. Augustine and St. Jerome.

Mention should be made here of the lexicographer Sextus Pompeius Festus, who lived and wrote sometime between AD 100 and 400. His *On the Meaning of Words* is an enormously important dictionary of archaic words and forms, containing a wealth of information also on older Roman legal and religious practice. It survives only in fragments, but we also have an abridged version of the whole work made by the eighth-century Lombard historian Paul the Deacon.

During this time the colloquial Latin spoken throughout the Empire, known as Vulgar Latin, was beginning to develop into the different dialects that would later become the Romance languages. See §§13.44ff. below.

Phonological developments of Latin

CONSONANTS

13.27. The main hallmark of Latin consonantism that sets it apart from its sister Italic languages, including the closely related Faliscan, is the outcome of the PIE voiced aspirates in word-internal position. In the other Italic dialects, these simply show up written as *f*. In Latin, that is the usual outcome word-initially, but word-internally the outcome is typically a voiced stop, as in *nebula* 'cloud' < **nebh-oleh*₂, *medius* 'middle' < **medh(i)jo-*, *angustus* 'narrow' < **anǵhos-*, and *ninguit* 'it snows' < **sni-n-gʷh-eti*. The details of these developments are left as an exercise at the end of this chapter.

13.28. Among the other changes to affect stops may be mentioned the loss of word-final *-d* after long vowels, as in the Classical Latin *o*-stem ablative sing. *-ō* from earlier *-ōd* (cp. Archaic *Gnaiuōd* 'from Gnaeus'). Also, as noted in §13.24, PIE **gʷ* became the glide *y*, as in *ueniō* 'I come' < **gʷem-jō*.

13.29. Consonant clusters. Among the many changes to consonant clusters, a few of special interest will be briefly mentioned. Voiced stops were lost before **j* or assimilated to it, as in the name of the god Jupiter, *Iūpiter*, stem *Iow-*, from **djey-* (Archaic genitive sing. *Dionos* 'of Jove') and in the comparative *maior* 'greater' (really *maiior*, with a geminate glide from the consonant cluster of earlier **magjos-*). The similar cluster **dy* became *b* at the beginning of a word, as in *bellum* 'war' (Archaic and poetic *duellum*), and *y* word-internally, as in *suāvis* 'sweet' < **suāduis*.

13.30. Rhotacism and other changes to *s*. Latin famously changed the sibilant *s* to *r* between vowels, a change known as *rhotacism*. Thus *mūs* 'mouse' has the plural *mūrēs* (< **mūs-ēs*), *genus* 'kind, race' (**genos*) has the plural *genera* (**genesa*), and the infinitive *-se* of *es-se* 'to be' appears as *-re* in vowel-stem verbs such as *amā-re* 'to love' and *duce-re* 'to lead'. This change happened during the historical period; early inscriptions still have intervocalic *s* (e.g. Archaic *iouesat* 'swears', Classical *iūrat*). Cicero noted in a letter that a certain Papirius Crassus officially changed the spelling of his name from Papisius in 339 BC, so the change probably happened not long before then.

13.31. In many other environments, especially before resonants, *s disappeared, as in the 'snow' words above. In clusters consisting of an s preceded by a liquid, it assimilated to the liquid. A famous example is *terra* 'land, earth' from **tersā*, root **ters-* 'dry'; originally the noun meant 'dry (land)'.

VOWELS

13.32. As stated above, Italic transformed the mobile accent system of PIE into a system characterized by stress on initial syllables. In Latin, this resulted in the weakening of vowels in non-initial syllables. The rules are rather complex, but in general terms a short vowel in an open syllable was weakened eventually to *i*. Thus compare *amicus* 'friend' with *mimicus* 'enemy', *legō* 'I choose' with *colligō* 'I collect', and *locus* 'place' with *ilicō* 'on the spot, right away'. In closed syllables, *a* was weakened to *e*, as in *affectus* 'affected' beside *factus* 'done', *ineptus* 'inept' beside *aptus* 'apt', while *o* was weakened to *u*, as in *onustus* 'burdensome' from **onos* 'burden' (Classical *onus*). In final syllables, the same rules usually apply, as in *artifex* 'craftsman' from **arti-fak-s* (from the bases *art-* 'skill' and *fac-* 'make, do') and *servus* 'slave' from earlier (Archaic) *seruos*; but before -s and -t, *e* weakened to *i*, as in *legis* 'you (sing.) lead' and *legit* 'he leads' from **leges* and **leget* (with the thematic vowel *-e-). In inscriptions from the Archaic period, many of these vowel weakenings had not yet happened.

13.33. Among some of the many other changes to vowels, two more may also be mentioned. An original *o* became *u* in most non-initial syllables: contrast Archaic Latin *malos* 'bad' (nomin. sing.) with Classical *malus*, and Archaic *servuom* 'slave' (accus. sing.) with *servum*; compare also *onustus* 'burdensome' < **onostos*, and the adverb *occultō* 'hidden', Archaic *oquoltōd*. An old *e* before nasals usually became raised to *i*, as in the preposition *in* 'in' from older *en* (preserved in inscriptions).

13.34. Long vowels and diphthongs. Long vowels were shortened in the late Archaic period in final syllables before any consonant except *s*. Thus contrast the 2nd sing. subjunctive *dūcās* '[that] you carry' with the 1st and 3rd singulars *dūcam* and *dūcat*.

13.35. Several of the old diphthongs were monophthongized to long vowels in Latin; those that survived were *ai* (spelled *ae* after the Archaic period), *au*, and in some cases, *oi* (spelled *oe*). The diphthong *ei* became *ī* (as in the dative ending of *patr-ī* 'for the father', cp. Archaic *Castorei* 'for Castor'), while *oi* became either *ī* (e.g. in the ablative pl. of *o*-stems, e.g. Classical *meīs sociūs* 'with my companions' but Archaic *meois sokiois*) or *ū* (as in *ūnus* 'one', Archaic accus. sing. *oino[m]*); it also sometimes remained unchanged, as in *foedus* 'treaty' and *moenia* 'walls'. PIE **eu* fell together with **ou* in Italic, and **ou* later became *ū*, as in *ūrere* 'to burn' (**h₁eus-e-*) and *iūmenta* 'teams (of oxen)' < Archaic *iouxmenta*.

13.36. Stress. The vowel weakenings in non-initial syllables discussed above indicate that the stress in early Latin was still on the first syllable, as in Italic (§13.11). In Classical Latin, however, the stress fell on the antepenult (third-to-last syllable) unless the following syllable was heavy. Thus *ánima* 'breath', *amābitur* 'he will be loved', and *adipiscímīni* 'you (pl.) are approaching' were all stressed on the antepenult, while *deárum* 'of goddesses', *amābúntur* 'they will be loved', and *adipíscor* 'I approach' were stressed on the penult. Latin borrowings into English are often stressed according to these rules.

Morphological developments of Latin

Latin morphology differs little from the picture outlined above for Common Italic. We may mention a few details.

NOUNS

13.37. Traditional Latin grammar divides nouns into five declensions. The first continues the PIE *ā*-stems and consists mostly of feminines. The second continues the *o*-stems, and consists mostly of masculines and neuters. The third continues the consonant stems as well as the *i*-stems, while the *u*-stems become the Latin fourth declension. The fifth declension is not an inherited type, but a medley of various formations that, due to sound change, all came to have a stem in *-ē*. The five declensions may all be illustrated by the paradigms given below of the nouns *terra* 'land', *lupus* 'wolf', *bellum* 'war' (neuter; given where different from *lupus*), *rēx* 'king' (and for *i*-stem forms, *turris* 'tower', where different), *currus* 'chariot' (and neuter *cornū* 'horn'), and *diēs* 'day'. Endings that are Archaic or poetic are given in parentheses (small capitals are for inscriptionally attested forms):

	I	II	III	IV	V
sg. N	<i>terr-a</i>	<i>lup-us</i> (-os), <i>bell-um</i> (-om)	<i>rēx</i>	<i>curr-us</i> , <i>corn-ū</i>	<i>di-ēs</i>
V	<i>terr-a</i>	<i>lup-e</i>	<i>rēx</i>	<i>curr-us</i> , <i>corn-ū</i>	<i>di-ēs</i>
G	<i>terr-ae</i> (-āī, -ās)	<i>lup-ī</i> (-osio)	<i>rēg-is</i> (-es, -os)	<i>curr-ūs</i> (-uos)	<i>di-ēī</i>
D	<i>terr-ae</i>	<i>lup-ō</i> (-ōī)	<i>rēg-ī</i> (-eī)	<i>curr-uī</i> , <i>corn-ū</i>	<i>di-ēī</i>
Ac	<i>terr-am</i>	<i>lup-um</i> (-om)	<i>rēg-em</i> , <i>turr-im</i>	<i>curr-um</i> , <i>corn-ū</i>	<i>di-em</i>
Ab	<i>terr-ā</i> (-ād)	<i>lup-ō</i> (-ōd)	<i>rēg-e</i> (-ed), <i>turr-ī</i> (-īd)	<i>curr-ū</i> (-ūd)	<i>di-ē</i>
pl. NV	<i>terr-ae</i>	<i>lup-ī</i> (-eī, -oe), <i>bell-a</i>	<i>rēg-ēs</i>	<i>curr-ūs</i> , <i>corn-ua</i>	<i>di-ēs</i>
G	<i>terr-ārum</i>	<i>lup-ōrum</i> (-ōrom, -um, -om)	<i>rēg-um</i> (-om), <i>turr-ium</i>	<i>curr-uum</i> (-uom)	<i>di-ērum</i>
D	<i>terr-īs</i>	<i>lup-is</i> (-ois)	<i>rēg-ibus</i>	<i>curr-ibus</i>	<i>di-ēbus</i>
Ac	<i>terr-ās</i>	<i>lup-ōs</i> , <i>bell-a</i>	<i>rēg-ēs</i> , <i>turr-īs</i>	<i>curr-ūs</i> , <i>corn-ua</i>	<i>di-ēs</i>
Ab	<i>terr-īs</i>	<i>lup-is</i> (-eis)	<i>rēg-ibus</i>	<i>curr-ibus</i> (-ibos)	<i>di-ēbus</i>

13.38. A few remarks may be appended. In the first declension, the genitive singular ending *-as* (as in the fixed phrase *pater familiās* 'head of a household') is the oldest; the Classical ending *-ae* (< earlier *-āī*) comes from the spread of the genitive *-ī* of the second declension (this ending also spread to the fifth, whence *-ēī*). In the second declension, the original *o*-stem nominative plural **-ōs* was replaced by the pronominal nomin. pl. in **-oi*, as happened in a number of other IE languages too (§6.53). Two genitive plurals in the *o*-stems are found, an older (in Classical times, poetic) one in *-um* (earlier *-om*), and a longer one in *-ōrum*, formed by analogy to *-ārum* in the first declension (itself from **-āsōm*). In the third declension, noteworthy is the presence of genitive singulars in **-es* (which became the standard *-is*) as well as **-os*

(as in Greek; limited to a few inscriptional attestations). The older ablatives all ended in *-d* (as per §13.22).

A locative case is still found vestigially in the first three declensions: *Rōm-ae* 'in Rome', *dom-ī* 'at home', *rūr-ī* 'in the country'.

VERBS

13.39. The Latin verbal system is the same as that described above for Italic (§§13.13ff.). The paradigms below will illustrate the active and passive forms in the present tense; the verbs are *amāre* 'to love' (first conjugation), *habēre* 'to have' (second), *dūcere* 'to lead' (third), *capere* 'to take' (third *-iō*), *audire* 'to hear' (fourth), and *esse* 'to be' (athematic). Archaic or poetic forms are indicated in parentheses and in the notes.

	I	II	III	III <i>-iō</i>	IV	Athem.
Active						
sg. 1	<i>am-ō</i> 'I love'	<i>hab-eō</i> 'I have'	<i>dūc-ō</i> 'I lead'	<i>cap-iō</i> 'I take'	<i>aud-iō</i> 'I hear'	<i>sum</i> ¹ 'I am'
2	<i>am-ās</i>	<i>hab-ēs</i>	<i>dūc-is</i>	<i>cap-is</i>	<i>aud-is</i>	<i>es</i>
3	<i>am-at</i>	<i>hab-et</i>	<i>dūc-it</i>	<i>cap-it</i>	<i>aud-it</i>	<i>est</i>
pl. 1	<i>am-āmus</i>	<i>hab-ēmus</i>	<i>dūc-imus</i>	<i>cap-imus</i>	<i>aud-imus</i>	<i>sumus</i>
2	<i>am-ātis</i>	<i>hab-ētis</i>	<i>dūc-itis</i>	<i>cap-itis</i>	<i>aud-itis</i>	<i>estis</i>
3	<i>am-ant</i>	<i>hab-ent</i>	<i>dūc-unt</i> ²	<i>cap-iunt</i>	<i>aud-iunt</i> ¹	<i>sunt</i>
Passive						
sg. 1	<i>am-or</i> 'I am loved'	<i>habeor</i> 'I am had'	<i>dūcor</i> 'I am led'	<i>capior</i> 'I am taken'	<i>audior</i> 'I am heard'	
2	<i>am-āris (-āre)</i>	<i>habēris (-ēre)</i>	<i>dūceris (-ere)</i>	<i>caperis (-ere)</i>	<i>audīris (-īre)</i>	
3	<i>amātur</i>	<i>habētur</i>	<i>dūcitur</i>	<i>capitur</i>	<i>auditur</i> ⁴	<i>itur</i> ⁵
pl. 1	<i>amāmur</i>	<i>habēmur</i>	<i>dūcimur</i>	<i>capimur</i>	<i>audimur</i>	
2	<i>amāmini</i>	<i>habēmini</i>	<i>dūcimini</i>	<i>capimini</i>	<i>audimini</i>	
3	<i>amantur</i>	<i>habentur</i>	<i>dūcuntur</i>	<i>capiuntur</i>	<i>audiuntur</i>	

¹ Archaic *esom*. ² Archaic *ueiunt* 'they live'. ³ Archaic *cōsentiant* 'they agree'. ⁴ Archaic *nancitor* 'he obtains'. ⁵ From *īre* 'to go', meaning 'one goes, people go'; archaic *itur*.

The verbs like *capere* are descended from **-je/o-* presents where the suffix **-je/o-* followed a light root (see §13.13); they take regular 3rd-conjugation endings except in the 1st person sing. and the 3rd pl. The 2nd sing. passive ending in *-re* is the older form, continuing **-se*; the regular Classical ending *-ris* is an example of double marking, since it ends with an added 2nd sing. active ending *-s*. The 3rd person endings in *-t* and *-nt* continue the PIE primary endings **-ti* and **-nti*.

THE PERFECT

13.40. Several of the types of perfect stems still resist straightforward historical explanation. Probably the most important of these is the so-called *v*-perfect, exemplified by such forms as *portāu-ī* 'I carried', *nēu-ī* 'I sewed', and *audīu-ī* 'I heard'. Some have tried to connect it with the *-u* in Sanskrit perfects like Ved. *dadāu* 'I/he gave'

(root *dā-*), *tastháu* 'I stood' (root *sthā-*); others have preferred to say that the pattern began with a verb whose perfect stem could have originally ended in **-V̄u-*, such as perhaps *gnōui* 'I knew' (**ġneh₃u-*). Not found in Sabellic (perhaps by accident) are perfects that continue *s*-aorists, such as *uēxī* 'I conveyed' (< **uēġh-s-*; §5.43).

13.41. The endings of the perfect are still recognizably descended from the PIE perfect endings, being conspicuously different from the regular active and passive endings, as illustrated by the paradigm of *gnōu-i* 'I have learned, I know' (archaic or poetic forms given in parentheses):

Singular	Plural
1 <i>gnōu-i</i> (-ei)	<i>gnōu-imus</i>
2 <i>gnōu-istī</i> (-isteī)	<i>gnōu-istis</i>
3 <i>gnōu-it</i> (-et, -eit)	<i>gnōu-ērunt</i> (-ēre, -erai)

One can see lurking under these the PIE endings **-h₂e* (> **-a*), **-(s)th₂e* (> **-sta*), **-e* in the singular, all extended by the primary active particle **-i* (§5.13), and an altered version of the 3rd plural ending **-ēr* (extended by the non-perfect 3rd pl. ending *-unt* in the familiar Classical ending *-ērunt*).

The later history of Latin

VULGAR LATIN

13.42. Most of the Latin that has come down to us consists of literature in an elevated style. A few works, such as the comedies of Plautus, the novel *Satyricon* of Petronius, and some letters of Cicero, contain colloquial language, and many inscriptions reflect the spoken idiom of the day. The spoken form of Latin, especially from about the third century AD on, is called *Vulgar Latin* (*vulgāris* 'pertaining to the common people') and forms the basis of the modern Romance languages.

Among the more extensive remains of early Vulgar Latin are the citations of "incorrect" forms collected in a work of the third or fourth century called the *Appendix Probi*, an appendix to a grammar by the grammarian Probus that was added by an unknown later hand. The *Appendix* consists of lists of words in their correct Classical spellings followed by their incorrect counterparts, such as *speculum non specum* (indicating that *speculum* and not *specum* is the correct spelling of the word for 'mirror'). The *Appendix*'s efforts were all in vain, for the "incorrect" spellings reflect pronunciations and forms that would all win out in the Romance languages. Thus for example the form *specum* (and not *speculum*) is the source of Italian *specchio* 'mirror'.

DEVELOPMENTS OF VULGAR LATIN

13.43. Unstressed vowels in internal syllables were syncopated (lost), as in *specum* just discussed. The distinction in length between long and short vowels was lost, and most diphthongs became monophthongized: the diphthong *oe*, for example, became *e*, as in Italian *pena* 'sorrow' from Lat. *poena* 'punishment'. In most areas

where Vulgar Latin was spoken, the velars became palatalized before front vowels, for example turning into [s] in French *cent* 'hundred' and [č] in Italian *cento* (both < Lat. *centum*).

There were also many changes in morphology. The number of cases in the noun was reduced, with prepositional phrases often taking over their functions: if one gave something 'to the king', one said *ad regem* rather than the dative *regi*. Eventually in the individual Romance languages, the cases were lost entirely except in personal pronouns. (The general development mirrors that from Old to Middle English; see §15.62.) Verbs did not go through as much formal reduction as nouns did, but periphrastic constructions (that is, constructions using "helping" verbs, as in English *I have seen*) became quite common. Typical of these were the compound perfect tense consisting of *habère* 'to have' plus the past participle (e.g. French *j'ai chanté* 'I have sung' < *ego habeo cantatum*) and the compound future tense consisting of an infinitive plus *habère* (e.g. French *chanterai* and Spanish *cantaré* 'I will sing' < *cantare habeo*).

The Romance languages

13.44. No texts in any of the vernacular descendants of Vulgar Latin survive from before the ninth century, and it is possible that none were written down before then. The language of administration and the Church was always Latin, and only when the vernaculars had diverged so much from Latin that ordinary people could no longer understand it did the need arise to use the vernaculars in writing and in the Church.

FRENCH

13.45. The earliest Romance language to be attested is French, a northern variety of which first appears in writing in the Strasbourg Oaths in or around the year 842. It is surely no accident that this is the first Romance language to have been written down, as it had diverged more strongly from Latin than the other varieties closer to Italy. Literary remains of French remain meager, however, until the twelfth century. The language is known as **Old French** until about 1403.

The part of Europe that is now called France had several varieties of Romance. In central France, especially around Paris, was spoken a variety that would become the dominant dialect already in the twelfth and thirteenth centuries and which developed into modern standard French. In the far north was **Norman French**, which spread to England following its conquest in 1066 by William the Bastard (as he is called in contemporaneous official documents, but now more familiar as "the Conqueror"). The Norman French that developed in England is often called **Anglo-Norman**, which flourished for at least two centuries before its eventual eclipse by Middle English in the fourteenth century. In the south were spoken varieties of French termed *langue d'oc* or **Occitan**, an important language of poetry in the Middle Ages until southern France was taken over by the north in the early 1200s. (The term *langue d'oc* refers to the way of saying 'yes' in this region, *oc* < Lat. *hoc* 'this'; the northern half of the country spoke *langue d'oïl*, their word for 'yes' being *oïl* [> modern *oui*] from Lat. *hoc ille* [*fecit*] 'he [did] this'.) Occitan varieties are still spoken in southern France,

the most prominent one being Provençal in the Provence and neighboring regions, attested from as early as the tenth century.

SPANISH

13.46. The Romance language spoken by the greatest number of people worldwide, Spanish, is first attested in the form of glosses on Latin texts dating probably to the mid-eleventh century. The earliest Spanish does not evince many dialect differences, but the famous twelfth- or early thirteenth-century epic poem *Cantar de mio Cid* (*Song of My Cid*, or *El Cid* for short) is in an early variety of Castilian, the dialect that would eventually become the standard.

As in France, the Romance varieties spoken in Spain are not homogeneous. In the northeast is Catalan, for centuries the official language of the kingdom of Aragon (until 1749). It is first attested in the twelfth century; before that, Catalan poets had written in Provençal. As this last fact attests, Catalan occupies an intermediate linguistic position between Spanish and the varieties of Occitan, although Spanish influence has grown over time. In the south, now-extinct Romance varieties collectively called Mozarabic were spoken; they were heavily influenced by the Arabic of Spain's Moorish invaders.

13.47. The northwestern dialects of Spain, especially Galician, are historically varieties of Portuguese, nowadays the second-most populous Romance language. Galician and Portuguese were a unitary language until the 1400s, called Gallego-Portuguese; scattered words in this language are recorded already in the late ninth century. The earliest true text dates to the late twelfth or early thirteenth century, and for two hundred years Gallego-Portuguese enjoyed a high literary prestige throughout most of the Iberian peninsula. Eventually Portuguese and Galician diverged; Portuguese formed its own literary standard beginning in the fifteenth and sixteenth centuries, soon eclipsing Galician on the European and the world stage. Brazilian Portuguese has itself diverged somewhat from European Portuguese in phonology and syntax.

ITALIAN

13.48. Italy has held a patchwork of dialects for most of the past millennium. The earliest records that can securely be called Italian are in the form of court records from the tenth century. The dialect of Florence became the basis of the standard literary language beginning in the thirteenth century; it was phonologically more conservative in several respects than other dialects, but admixtures of forms from them have continued to shape the standard.

SARDINIAN

13.49. The island of Sardinia, a Carthaginian colony before it became Roman territory soon after Carthage's defeat in the First Punic War (264–241 BC), is where Sardinian is spoken today. Sardinian is attested quite early, near the end of the eleventh century, but little literature has ever been written in it. Central dialects are remarkably conservative in one famous respect: the velars did not become palatalized before front vowels (§13.43), as in their word for 'hundred', *kentu* (Lat. *centum*).

ROMANIAN

13.50. Romanian (or Rumanian), spoken in Romania, Moldova, and neighboring areas, descends from the Latin of the Roman provinces of Dacia and Illyricum. Beginning in the third century, the region fell out of Roman control and was taken over successively by Goths, Bulgaria, Hungary, the Ottoman Empire, and Russia. The linguistic influence especially of Slavic and Hungarian was far-reaching. The first Romanian text dates only from 1521; the Cyrillic alphabet was used until 1859. All of the numerous divergent dialects of Romanian are nearly extinct except for the standard, called Daco-Romanian.

RHAETO-ROMANCE

13.51. Certain minor and isolated Romance languages spoken in Switzerland and northern Italy comprise the Rhaeto-Romance or Rhaetian group. Most prominent are Romansh, Engadine, and Ladin.

Archaic Latin text sample A

13.52. The so-called Duenos inscription, found in Rome, inscribed around a three-part terracotta bowl and dating to c. 500 BC. Only the first and third lines can be understood. The original has no word-divisions. Latin inscriptions are collected in the multi-volume *Corpus Inscriptionum Latinarum* (CIL): this one is CIL I² 4.

iouesatdeiuosqoimedmitatneitedendocosmisuircosied
 astednoisiopetoitesiaipacariuios
 duenosmedfecedenmanomeinomuenuoinemedmalostatod

The first and third lines with word-divisions added read as follows:

iouesat deiuos qoi med mitat nei ted endo cosmis uirco sied
 duenos med feced en manom einom duenoi ne med malos tatod

He who sends (?) me swears by the gods: if a girl is not nice to you, [...] A good (man) made me for "good going" for a good (man). Let no bad (man) steal me.

13.52a. Notes. 1. *iouesat*: Classical *iūrat*, 'swears'; intervocalic *s* is still unrhottacized, and contraction of the sequence *-oue-* to *-ū-* has not happened yet either. *deiuos*: 'gods', accus. pl., Classical *diuōs*. *qoi*: '(he) who', Classical *qui*, relative pronoun. *med*: *med*, 'me', Archaic accus. sing.; Classical *me*. The *-d* is secondary and of uncertain origin, but may have spread from the ablative (also *mēd*, ultimately from PIE **med*). *mitat*: 3rd sing. of an unknown verb, probably meaning 'gives' or 'sends' or the like. It most resembles the Classical Latin present subjunctive *mittat* of the verb *mittere* 'to send'. *nei*: 'if not', Classical *ni*. *ted endo*: 'toward you', with postposed preposition *endo*, Archaic for *in*. The form *ted* is parallel to *mēd* above. *cosmis*: 'nice', Classical *cōmis*, with the cluster *-sm-* still preserved. *uirco*: 'girl', Classical *uirgo*. The letter *C* was originally used for both *c* and *g* (*G* is simply an alteration of *C*); cp. the abbreviation *C.* for the name *Gaius*. *sied*: 3rd sing. present subjunctive of *esse* 'to be', with *-d* from PIE secondary **-t*.

2. *duenos*: 'good', Classical *bonus* (*du* regularly became *b-*); used as a noun. The dative *duenōi* a few words later preserves the long diphthong *-ōi* (Classical *bonō*). *feced*: 'made', 3rd sing. perfect; Classical *fēcit*. *en manom cinom*: of disputed interpretation; perhaps 'for good going'. Under this interpretation, *en* is the preposition *in* 'in'; *manom* is from an otherwise unattested adjective that is the base of the adverb *māne* 'tomorrow' (cp. French *de bonne heure* 'tomorrow', lit. 'of a good hour'); and *cinom* is a verbal abstract noun from the root **h₁et-* 'go'. *ne; nē*, negative with the imperative at the end of the line. *malos*: 'a bad (man)', Classical *malus*. *tatod*: 3rd sing. imperative of an otherwise unattested verb **tāre* 'to steal', cp. Hitt. *tāēzzu* 'steals'.

Archaic Latin text sample B

13.53. Excerpt from the so-called *suovitaurlia* prayer, recorded by Marcus Porcius Cato (Cato the Elder), *Dē Agrī Cultūrā* (On Agriculture) 141.1ff. This was a prayer to Mars on the occasion of the purification of a field, during which a sow (*sūs*), sheep (*ovīs*), and bull (*taurus*) were sacrificed. The prayer has phraseology that is in places identical to that seen in the selection from the Umbrian Iguvine Tables below, and which has further analogues in Indo-Iranian (see the Notes). The repetitive language is typical of sacral poetic style; note also the alliterating pairs as in the Umbrian and South Picene texts further below (*uiduertātem uastitūdinemque, pāstōrēs pecuaque*, etc.).

Mārs pater tē precor quaesōque uti siēs uolēns propitiū mihi domō familiaeque nostrae, [...] uti tū morbōs uisōs inuisōsque uiduertātem uastitūdinemque calamitātem intemperiasque dēfendās prohibessis āuerruncāsque utique tū frūgēs frūmenta uinēta uirgultaque grandire beneque ēuenire sīs pāstōrēs pecuaque salua seruassis dūisque bonam salūtem ualētūdinemque mihi domō familiaeque nostrae [...]

Father Mars, I beseech and entreat you, that you be willing (and) propitious to me, (my) house and our household, [...] that you ward off diseases seen and unseen, banish barrenness and devastation, and sweep away destruction and bad weather, and that you allow the fruits and grains, the vineyards and shrubbery to grow tall and come out well, that you protect shepherds and livestock, and that you give good safety and health to me, (my) home and our household [...]

13.53a. Notes. *Mārs pater*: also *Mārspiter*, cp. *Itā-piter* and Gk. *Zeū páter*. *precor*: 'I beseech', a deponent (§13.20) 1st sing. present, from PIE **prek-* 'ask' (which also formed a *ske*-present **prk-ske-* in Lat. *poscō* 'I demand'). *-que*: 'and', cognate with Gk. *te*, Skt. *ca*; §7.26. *uti*: 'that', conjunction. *siēs*: '(that) you be', Archaic subjunctive, Classical Latin *sīs*, a continuation of the IE optative **h₁s-ieh₁-s* (§13.19); cp. Ved. *syās*. *uolēns*: 'willing', present participle of *uolō* 'I want', an athematic verb. *mihi*: 'to me', dat. sing., probably from earlier **mehei* < **megh(e)i*, cp. Ved. dat. *māhy-am*. *domō*: 'to (my) home', dat. sing. of *domus* 'house'; PIE root **dem-*. *familia*: 'household', dat. sing., derived from *famulus* 'household slave'. *morbōs*: 'diseases', o-stem accus. pl.; the ending was earlier **-ōns* (§6.55). *uisōs*: 'seen', not directly from expected **uid-to-* (which should have given **uisso-*) but apparently remade to the full-grade **ueid-to-*. *uiduertātem*: only occurs here in all of Latin literature and literally means 'widowhood', built to the word *uidua* 'widow' (PIE **uidheueh₂*, the source of Eng. *widow*) and probably modeled on *libertāt-* 'freedom', which is sort of a semantic opposite. *dēfendās prohibessis āuerruncās*: '(that) you ward off, banish, sweep away'; these verbs are not entirely synonymous, as each goes with a particular evil. The forms *dēfendās* and *āuerruncās* are ordinary

ā-subjunctives (§13.19), while *prohībessīs* is an Archaic Latin subjunctive containing the PIE optative morpheme (*-ī-* < **-ih₁-*) added to an *-s-* or *-ss-* of uncertain origin. *grandīre*: 'to grow tall', a denominative from the *i*-stem adjective *grandis* 'tall'. *bene*: 'well'; the earlier form was *duene*, and if we restore that here we get *dueneque ēuenīre sirīs*, with a lovely succession of long and short *e*'s and *i*'s. *ēuenīre*: 'to come out', a compound of *uenīre* 'come', PIE **g^{em}-* (> Goth. *qiman*, Eng. *come*, Ved. root *gam-*). *sirīs*: '(that) you allow'. *pāstōrēs pecuaque*: 'shepherds and livestock'; the phrase signifies movable wealth in the form of slaves (which shepherds were considered in archaic Roman society) and livestock, or two-footed and four-footed wealth. Recall §2.10. Both *pāstōrēs* and *pecua* are good PIE inheritances: **peh₂-* 'protect' (Ved. *pā-*, with *-s-* also in Hitt. *paḥš-* with the laryngeal preserved) and **peku-* (Ved. *pāsu* 'cattle', German *Vieh* 'animal, cattle', Eng. *fee*). *salua seruassīs*: '(that) you protect, keep safe', an inherited Italic sacral formula, seen also below in Umbrian *salua seritu. dūis*: '(that) you give', an Archaic subjunctive of *dare* 'give'; of disputed origin. *bonam*: 'good'; earlier **duenam* (see the *Duenos* inscription above).

Faliscan

13.54. Faliscan was spoken by a people known in ancient times as the Falisci, in and around the city of Falerii (nowadays *Civita Castellana*), about 60 km north of Rome in southern Etruria. It is known from about 300 inscriptions from the seventh to the second centuries BC. With a few notable exceptions, the inscriptions are mostly quite short, and our knowledge of the language correspondingly scanty. The Falisci were closely related to the Latini, and their language was quite close to Latin as well, as for example in its treatment of the labiovelars, which were preserved as such and not changed into labial stops (§13.24). It does not pattern with Latin in all matters of phonology, however: unlike Latin and like Sabellic, for example, the voiced aspirates became *f* word-internally, as in *pipafo* 'I will drink' (with *-f-* from **-bh-*) and *efiles* 'officials, aediles' (with *-f-* from **-dh-*). In word-initial position, the voiced aspirates often became *h-*, as in *huticilom* 'a little cash' in the text sample below (with *h-* from **gh-*).

Faliscan text sample

13.55. The Ceres inscription, Ve 241. (Inscriptions in Faliscan and other Italic dialects are conventionally cited by their catalogue number in the collection assembled by Emil Vetter, *Handbuch der italischen Dialekte* [*Handbook of the Italic Dialects*], Heidelberg, 1953.) Found at Falerii; c. 600 BC. The text is given as it appears in the original; the colons represent interpuncts, dot-like punctuation marks used in ancient inscriptions to separate words or phrases.

ccres : farme[la]tom : louff[i]rui[no]m : []rad
 euios : mamaxextosmedf[i]f[i]qod :
 prauiosurnam : sociaipordedkarai :
 eqournela[t]itelafitaidupes :
 arcntelomhuticilom : pe : parai[]douiad

Let (?) Ceres [. . .] ground grain, Bacchus wine.
 Mama (and?) Zextos Euios fashioned me.

Prauios gave the urn to his dear girlfriend.
 I, a little *titela*-urn, . . .
 have given birth to a little money. May it give (?).

13.55a. Notes. *ceres*: 'Ceres', the goddess of grain; from PIE **keres* (cp. Hitt. *karaš* 'grain'). *far*: 'grain', Lat. *far*, Eng. *bar-ley*. *me[la]tom*: 'ground', if correctly restored (= Lat. *molitum*, PIE **melh₂*- 'grund'); 'ground' is a frequent epithet of words for 'grain' in Italic, especially in religious contexts. *louf[i]r*: 'Bacchus', the Faliscan form of the native Italic word for the god of wine; cp. Lat. *Liber*, with Lat. *-b-* and Fal. *-f-* from **-dh-*. *[i]rad*: 3rd sing. of some verb, probably subjunctive; 'let . . .' or 'may . . .'. *euios mamazextos*: probably Euios is the family name going with two first names, Mama and Zextos; but uncertain. *med*: 'me', cp. Archaic Lat. *mēd*; see §13.52a above. *f[i]f[i]qod*: 'have fashioned', most likely a 3rd pl. like Lat. *-unt* (Archaic Lat. *-ont*), without the nasal written, and from secondary **-nt* rather than primary *-nti* (where the *-t-* would have remained). The *-q-* stands for *-g-* (compare *ego* for *egō* two lines below), and the form seems to be a reduplicated perfect of the Faliscan equivalent of Lat. *fiŋgō* 'fashion, create' (PIE **dheŋh-*, Eng. *dough*). *sociai*: 'girlfriend', dat. sing., Lat. *sociae*, ultimately from PIE **sekʷ-* 'follow, accompany'. The separation of the noun from its modifier *karai* 'dear' by an intervening verb is a common poetic stylistic device in the older IE languages. *porded*: 'gave', 3rd sing. *karai*: 'dear', dat. sing. fem., Lat. *cārae*; from PIE **keh₂-* 'love, desire' (> Ved. *kāmas* 'love', Eng. *whore* [prehistorically 'girlfriend' or 'beloved']). *urnela*: 'little urn', a diminutive like the following *[ti]tela*. The remainder of the line is not understood. *arcentelom*: 'little bit of money', cp. Lat. *argentum* 'silver'. *hucilom*: basically means the same thing as *arcentelom*; *hut-* may be the Faliscan cognate of the Latin root *fud-* 'pour' (nasal-infix present *fundō* 'I pour'); therefore *hucilom* is literally 'a little pourable stuff', i.e. liquid assets (same metaphor as in English!). The IE root is **gheu-* (Gk. *khé(w)ō* 'I pour', Eng. *in-got*). *pe*: *parai*: 'I have given birth to', reduplicated perfect 1st sing. with the archaic ending *-ai* < **-h₂e* plus the *hic et nunc* particle *-i* (§5.13; cp. also §13.41); cp. Lat. *pariō* 'I give birth' and *parēns* 'parent'. Why an interpunct separates the reduplicating syllable from the rest of the word is uncertain. *[doui]ad*: maybe a 3rd sing. subjunctive of the Faliscan cognate of Lat. *dō* 'I give', reminiscent of the Archaic Lat. subjunctive *duit* (see §13.53a).

Sabellic (Osco-Umbrian)

13.56. The Sabellic languages derive their name from the Sabelli, another name for the Samnites; both names are in turn etymologically connected with the name of the Sabines. The Sabellic-speaking peoples originally inhabited an area to the east and northeast of Latium, but a subgroup of them that spoke Oscan migrated southward into Campania around the middle of the first millennium BC. After this migration, Sabellic languages were spread out over most of Italy except along the western coastal strip; centuries later they would be eclipsed by Latin and would die out.

Common Sabellic developments

13.57. Superficially, the Sabellic languages, especially Umbrian, look strikingly different from Latin because of the far-reaching effects of a few basic sound changes. Sabellic in general also presents considerable divergences from Latin in nominal and verbal morphology. However, peeling away these differences one finds that the Sabellic languages really behave much the same as Latin.

Sabellic phonology

13.58. Due first of all to the limited number of inscriptions we have in Sabellic, and second to their inconsistent spelling, any sketch of Sabellic phonology, both historical and synchronic, must be tentative. The following conclusions are among the most secure.

The voiced aspirates became *f* not only word-initially (as usually in Latin) but also word-internally: Osc. *prúfatt* 'he approved' (cp. Lat. *probāuit*, from **probh-*), *meñai* 'in the middle' (Lat. *media-* < **medh(i)ā-*), and Umbr. *vufu* 'votive' (< **uog^h-ro-*, cp. Lat. *uou-ēre* 'to vow'). As sometimes in Latin, **gh* becomes *h* word-initially: Osc. *húrz* 'yard' (cp. Lat. *hortus* < **ghort-*).

Also characteristic of Sabellic, as mentioned in §13.24, is the development of the labiovelars to labial stops: Osc. fem. accus. sing. *paam* 'whom' (Lat. *quam*), Osc. *biítam* 'life' (Lat. *uítam*) < **g^ui^h-*. The syllabic nasals may have developed differently in Sabellic vis-à-vis Latino-Faliscan, as discussed in §13.8.

13.59. Two important developments of the vowel system may also be mentioned. First is the syncope (loss) of short vowels in final closed syllables, as illustrated by Umbr. accus. pl. *manf* 'hands' (< **manuf* < **manuns*, cp. Lat. *manūs*) and Osc. nomin. sing. *húrz* 'yard' (pronounced *hurts*, from **hortos*). Second is the change of final **-ā* to *-ū*: Osc. *viú* 'road' (**uiā*, cp. Lat. *uia*), Umbr. *mutu* 'punishment' (**moltā*), *toto* 'tribe' (**toutā*).

Morphology

13.60. The Sabellic languages preserve more distinctions between primary and secondary verbal endings in the third person than Latin. The active plural primary ending is *-nt* (as in Osc. present indicative *stahint* 'they stand'), while the secondary ending is *-ns* (as in Osc. imperfect subjunctive *patensíns* 'they would open'). Umbrian has a further distinction, lost elsewhere in Sabellic, in the passive (see §13.70).

13.61. Not found in Latino-Faliscan is an infinitive in *-fi-* from PIE **-dhj-* (§5.54), which can have active or passive meaning: Osc. *sakrafír* 'to consecrate, be consecrated' and Umbr. *pihafi* 'to be offered'.

13.62. Umbrian, together with Oscan, has several peculiarities in the morphology of the perfect system not shared with Latin. Only Sabellic has an *f*-perfect (Osc. *fufens* 'they were' from the verbal stem *fu-* 'be') and a *tt*-perfect (found only in Oscan, as in *prúfatt* 'approved'). Their origins are disputed.

13.63. The system of nominal inflection is quite similar to that of Latin with the exception of some endings. The *o*-stem genitive singular has been replaced by the *i*-stem ending *-eis*: Osc. *sakarakeís* 'of the temple', Umbr. *popler* 'of the people' (< **popleis*, contrast Lat. *populí*). Unlike Latin, Sabellic preserves the inherited *o*-stem nominative pl. ending **-ōs*: Osc. *Núvlanús* 'the inhabitants of Nola', South Picene *safínús* 'Sabines'. And judging from Umbrian, it also preserves the inherited distinction between the nominative and vocative singular of feminine *ā*-stems, which were *-ā* and **-a*, respectively: contrast Umbrian nominative *mutu* 'punishment' with the vocative *Tursa* (a proper name).

Umbrian

13.64. Umbrian is almost entirely known from the seven Iguvine Tables, discovered in 1444 in the Italian town of Gubbio (in classical times Iguvium), about 160 km north of Rome near Perugia. It is said that the original number of tablets was nine and that two were later lost after their discovery, although there is evidence that this is erroneous. The Iguvine Tables constitute the longest text in any non-Latin language of Italy, containing a set of ritual instructions for a class of priests called the Atridian Brethren and containing about 4000 words in all. The content is immensely valuable for our knowledge of native Italic religion and cultic practice, about which the Romans themselves did not tell us much (as a result of the Hellenization of Roman religion). Not all the tablets were inscribed at the same time: tables V and VIb are linguistically more recent and easier to understand. Tables VI and VII are in the Latin alphabet and were probably written down shortly after the Social War (90–89 BC) but copied from an older original. The rest are in the native Umbrian alphabet and are conventionally transliterated in boldface.

Umbrian is otherwise known from a few dozen scattered inscriptions from the sixth to the first centuries BC. Already by the time of its earliest attestation the language had undergone many of the phonological reductions that make it look more unlike Latin than any other Italic dialect. Some of these are listed below.

Consonants

13.65. As in Latin, *s* was rhotacized between vowels, e.g. *erom* 'to be' (**esom*), but in later Umbrian at the ends of words too, as in the phrase *pre ueireir treblaneir* 'before the Trebulan gate' (written *preueres treplanes* in the older portions of the Iguvine Tables), with the ablative pl. ending *-eir* < **-eis*. The stop *d* became spirantized between vowels and sometimes before consonants to a voiced fricative sound written *rs* in the Latin alphabet and with a modified *r*, transcribed *ř*, in the native alphabet, as in *dirsa* (*teřa*) 'he should give' (**didāt*) and *arsfertur* (*ařfertur*) 'priest' (< **ad-fer-tōr*).

13.66. Characteristic of Umbrian (and South Picene) is the unusual change of word-final **-ns* to *-f*, as in accusative plurals (e.g. *nerf* 'magistrates' < **ner-ens*, vitlaf 'yearlings' < **vitelāns*) and in the present participle *zeřef* 'sitting' < **sedens*.

13.67. Various consonant-cluster simplifications further altered the look of the language. They can be exemplified by such forms as *ape* 'when' < **atpe* < **atkwe* (cp. Lat. *atque* 'and'); *une* 'water' (ablative) < **udni*; *uze* '(on the) shoulder' < **omse* < **omesei* (cp. Lat. *umerus*); *testru* 'to the right' < **dekstero-* (cp. Lat. *dexter*); and *kumatir* 'crumbled' < **komm(al)a*teis.

13.68. Later Umbrian palatalized *k* before front vowels or **i* to a sound written with a letter transliterated as *ç* (native alphabet) or *š* (Latin alphabet): *façia* 'let him make' < **fak-jad*; *šihitu* 'girded' (**kink-to-*); *çerfer* 'for Ceres' (**kerezeis*).

Vowels

13.69. Short vowels in many interior syllables, especially before single consonants, were lost, as in *actud* 'let him do' (**agetōd*, cp. Lat. *agitō*) and *vitlaf* 'yearlings' above (**vitelāns*). Whereas Oscan retains the diphthongs intact, Umbrian monophthongized

them: *ocrer* 'of the mount' < **ocreis*; *prephobotatu* 'let him trample' < **prai-plautâtōd* (with -*oho-* rendering *ō*); *ueres* 'gates' (abl. pl.) < *(*d*)*uerois*; and *muneklu* 'little gift' < **moi-ni-tlom*.

Morphology

13.70. Noteworthy in Umbrian is a distinction in the 3rd person mediopassive endings between *-ter*, found in present indicatives like *herter* 'there is need', and *-tur*, found in the subjunctive, as in *emantur* 'let them take'.

Umbrian text sample

13.71. From the Iguvine Tables, tablet VIa, lines 27–31. This section contains a prayer to Jupiter Grabovius, an Umbrian deity. The dots are interpuncts.

(27) . . . dei . crabouie . persei . tuer . perscler . uaseto . est . pesetomest . peretomest
(28) frosetomest . daetomest . tuer . perscler . uirseto . auirseto . uas . est . di .
grabouie . persei . mersei . esu . bue (29) peracrei . pihaclu . pihafei . di . grabouie
 . pihatu . ocre . fisi . pihatu . tota . iouina . di . grabouie . pihatu . ocrer (30) fisier
 . totar . iouinar . nome . nerf . arsmo . ueiro pequo . castruo . fri . pihatu . futu .
 fos . pacer . pase . tua . ocre fisi (31) tote . iouine . erer . nomne . erar . nomne .
 di . grabouie . saluo . seritu . ocre . fisi . salua . seritu . tota . iouina .

(27) . . . Jupiter Grabovius, if in your sacrifice (anything) has been done wrongly, mistaken, transgressed, (28) deceived, left out, (if) in your ritual there is a seen or unseen flaw, Jupiter Grabovius, if it be right for this (29) yearling ox as purificatory offering to be purified, Jupiter Grabovius, purify the Fisian Mount, purify the Iguvine state. Jupiter Grabovius, purify the name of the Fisian Mount (and) of the Iguvine state, purify the magistrates (and) formulations, men (and) cattle, heads (of grain) (and) fruits. Be favorable (and) propitious in your peace to the Fisian Mount, (31) to the Iguvine state, to the name of that, to the name of this. Jupiter Grabovius, keep safe the Fisian Mount, keep safe the Iguvine state.

13.71a. Notes. 27. dei crabouie: 'Jupiter Grabovius', voc. sing. It has been suggested, speculatively but intriguingly, that *Grabovius* is from Illyrian (presumably via Messapic, cp. §20.21) and means 'of the oak', cp. Illyrian *grabion* 'oak wood', Polish *grabowy* 'made of white beechwood', Modern Greek (Epirotic) *grabos* 'type of oak'. Oak in various ancient IE societies is mythologically associated with the god of thunder (§2.21). persei: 'if', also *pīrsi*; apparently from **k^wid-id*, cp. Lat. *quid* 'what'; see §13.65 on *rs* from intervocalic **d*. tuer perscler: 'of your sacrifice', genit. sing., with rhotacism of the original final *-s. uaseto est: 'there has been a fault'; should be *uasetomest* like the following with the neuter sing. -*m* intact. pesetomest: 'has been mistaken', cognate with Lat. *peccatum est* 'it has been sinned', but with stem vowel -*ē*- rather than -*ā*-. peretomest: 'has been transgressed', lit. 'gone through' (**per-ei*).

28. frosetomest: 'has been deceived'; *fros-* seems equatable with Lat. *fraud-* 'fraud'. daetomest: 'has been left out'. uirseto: 'seen' < **uidētom*, a different formation from Lat. *uīsum* 'seen' (**ueid-tom*). For the phrase 'seen (or) unseen fault', compare Lat. *morbōs uīsōs inuīsōsque* 'seen and unseen diseases' in the *suovitaurlia* prayer above (§13.53). mersei: *mers sei*, '(if) it be right'. The first word is from **medos*, from the root **med-* 'to take appropriate measures', and *sei* is the 3rd sing. subjunctive of 'to be', equivalent to Lat. *sit*. esu bue: 'with this ox', abl. sing.

29. *peracrei*: 'yearling', abl. sing.; dissimilated from *peraknei*, its form in older parts of the Tables. The root *-akn-* 'year' is dissimilated from the same **atn-* that became Lat. *annus* 'year'. *pihaclu*: 'offering', Lat. *piaculum*. *pihafei*: 'to be purified', passive infinitive in *-f(e)* (§13.61). *pihatu*: 'purify', 3rd sing. imperative but used like a 2nd sing. *ocre fisei*: 'the Fisian mount', accus. sing., again without writing final *-m*. *tota*: 'people, state', accus. sing.; cp. Gaulish *Teutatis* 'god of the people'. *iouina*: 'Iguvine'.

30–31. *nome*: 'name', accus. sing., cp. Lat. *nōmen*. *nerf*: 'magistrates', accus. pl., from IE **h₂ner-* 'man, hero'. The word is common to all Sabellic languages (including South Picene), but is not found in Latin except in the personal name *Nerō*, originally 'having manly strength' or the like. *arismo*: 'formulations', accus. pl. *ueiro pequo*: 'men and livestock', a phrase inherited from PIE; it occurs in Indo-Iranian as a term for movable wealth, as in Av. *paṣu vīra*. *castruo*: 'heads' (of grain), neut. accus. pl. of a *u*-stem < **kastruā*. *fri*: 'fruits', accus. pl. (final *-f* not written), cp. *frū-* in Latin *fructus* 'fruit'. *futu*: 'let there be', 3rd sing. imperative of *fu-* 'be' (in Lat. *fu-tūrus* 'about to be'; PIE **bhūH-*, also in Eng. *be*). *pase*: 'in peace', abl. *saluo*: 'safe, whole', Lat. *saluus*. *seritu*: 'hold, keep, preserve', 3rd sing. imperative. The phrase *saluo seritu* is the same as Lat. *salua seruassis* in the *suovitaurla* prayer above.

South Picene

13.72. South Picene is known from nearly two dozen inscriptions from an area in east-central Italy called Picenum in ancient times; the inhabitants were called the Piceni. Although South Picene inscriptions have been known for some time, due to difficulties posed by the script they remained essentially a closed book until very recently. In the 1980s it was finally realized that two symbols (. and :) that had always been assumed to be interpuncts were instead the letters *o* and *f*; needless to say, this discovery dramatically improved our ability to read these texts, and further advances in interpretation have been continuing apace. Our South Picene documents date from the beginning of the sixth to the third century BC; the earliest ones are among our oldest preserved texts in any Sabellic language.

South Picene appears to be more closely related to Umbrian than to Oscan. It is unrelated to another language of the region called North Picene, a non-IE language preserved in a single unintelligible text.

South Picene text sample

13.73. TE. 2, a gravestone found in Bellante near Teramo, south of Piceno. The inscription is poetry in the archaic Italic strophic style (see §§2.42ff.); except for the first word it consists of alliterative word-pairs (*viam videtas, tetis tokam*, etc.).

postin : viam : videtas : tetis : tokam : alies : esmen : vepses : vepeten

Along the road you see the "roga" of Titus Alius (?) buried (?) in this tomb.

13.73a. Notes. *postin*: 'along', Umbrian *pustin*. *videtas*: probably 'you see', 2nd pl., equivalent to Lat. *videtis* 'you (pl.) see'; passers-by are the addressees. *tetis alies*: apparently the name, in the genitive, of the man buried there. *tokam*: cognate with Lat. *togam* (accus. sing.), but the exact sense is uncertain ('covering'? The root is **(s)teg-* 'cover'). In many early Italic inscriptions *k* or *c* was used for voiced *g*. *esmen*: locative of the demonstrative stem *e-*, cp. Umbr. *esme*; morphologically identical to Sanskrit *asmīn* 'in this'; see §7.8. *vepses*: perhaps 'buried'; unclear. It might be a past participle of the sort seen in Lat. *lapsus* 'slipped'. *vepeten*: perhaps 'tomb', with locative in *-en*.

Oscan

13.74. Oscan is known from close to 400 mostly short inscriptions from central and southern Italy dating primarily from the fourth century BC into the first century AD. There are far more Oscan inscriptions than Umbrian ones, but they are mostly quite short and so our knowledge of the language is less secure.

13.75. Oscan is on the whole more conservative than Latin in its vowels, and more conservative than Umbrian in many other phonological respects. It has suffered very little phonetic reduction since Common Sabellic times. It even underwent at least one change going in the opposite direction – the epenthesis (insertion) of a vowel to break up consonant clusters consisting of a resonant and another consonant, as in *aragetud* ‘with money’ (cp. Lat. *argentō*) and *sakarater* ‘it is consecrated’ (cp. Lat. *sacrātur*).

13.76. The alphabets used to write Oscan do not distinguish the differences in the vowels equally well. In the north, where the “North Oscan” idioms of Paelignian, Marrucinian, and the other languages enumerated below in §13.78 were spoken, the Latin alphabet was used, which distinguished five vowels (*i e a o u*). In the central part of Oscan territory (Campania), where the Oscans were in contact with the Etruscans, they used the Etruscan alphabet; this is often referred to as the native Oscan alphabet, and is transliterated in boldface. In the south (as in Bruttium and Lucania), the Greek alphabet was used in the early inscriptions, and later the Latin alphabet.

Oscan text sample

13.77. Ve 11, from Pompeii; first century BC. Some expressions are translations of Latin administrative phraseology. Note that some words are split up across the end of a line.

v . aadirans . v . eitiuvam . paam
 vereiiaī . pūmpaiianai . trīstaa
 mentud . deded . eīsak . eitiuvad
 v . viinikiis . mr . kvaisstur . pūmp
 aīans . trībūm . ekak . kūmben
 nicīs . tancimud . ūpsannam
 deded . isīdum . prūfatted

The money that V(ibus) Atranus, (son) of V(ibus), gave to the Pompeian gatekeepers in his will – with that money V(ibus) Vinicius, (son) of M(aras), the Pompeian quaestor, gave, with agreement of the assembly, (for) this house to be built. The same one approved (it).

13.77a. Notes (selective). *eitiuvam*: ‘money’, lit. ‘stuff that goes (in)’, from *ei-* ‘go’. It is from earlier **eituvam*; the *-t-* became *-ti-* before the *-u-*, as in varieties of English where *tune* is pronounced *tyune*. *paam*: ‘which’, PIE **kʷām*; relative adjective modifying *eitiuvam*, so literally “Which money V. A. gave . . .” Both *eitiuvam* and the subject of the relative clause appear before the relativizer; recall §8.26. *vereiiaī*: ‘gatekeepers’, dat. sing. of an abstract noun referring to a class of youths having some connection with gates. *pūmpaiianai*: ‘Pompeian’; the word is derived from the numeral ‘five’, Osc. *pūmp*. *trīstaa* *mentud*: ‘by testimony, in the will’, abl. sing. of the Osc. cognate of Lat. *testāmentum*, both from **tri-st-* ‘stand by as the third’, i.e. to witness; the sequence **tris-* became **ters-* and then *tes-* in Latin by a sound

change. *deded*: 'gave', 3rd sing. perfect (= Lat. *dedit*, Archaic Lat. *dedet*); the old perfect ending **-e(i)* has been replaced with the aorist **-et*. *eisak*: 'with that', fem. abl. sing., from **eisād-k*, the *-k* being cognate with the particle *-c* in such Latin demonstratives as *hic* 'this', *nunc* 'now', etc. The accus. *ekak* below is from **ekām-k* with loss of the nasal. *kvaisstur*: 'quaestor', an official charged with tax-collection and other financial duties. The Oscan word is borrowed from Lat. *quaestor*. *trībúm*: 'house', from **trēbom*, from a root **treb-* found in Eng. *thorpe*. *kūmbennieis*: 'of the senate', genit. sing.; a noun formed from *kum-* 'together' (Lat. *con-*, *com-*) and *ben-* 'come' (PIE **g^{em}-*). *tancinud*: 'by decision', abl. sing. *-ud* < **-od*. Phonetically this begins *tang-*, from the same root as Eng. *think*. The Oscan phrase *kūmbennieis tancinud* is a loan-translation of the Latin bureaucratic phrase *senātūs sententia*. *úpsannam*: 'to be built'; *-annam* is equivalent to the Latin gerundive (future passive participle) in *-andam*. The Oscan root *úps-* is cognate with Lat. *opus* 'work'. The Oscan phrase *úpsannam deded* is a loan-translation of Lat. *faciendam cūrāuit* 'saw to it that . . . be done'. *isidum*: 'the same one', equivalent to Lat. *is* 'he, that one' + *īdem* 'the same (one)'. *prúfatted*: 'approved', an Oscan *tt*-perfect (§13.62), from *prúfa-* = Lat. *probā-*.

Other Sabellic Languages

13.78. Besides Oscan, Umbrian, and South Picene, scattered inscriptions in over half a dozen other Sabellic languages from central Italy east of Rome are known, collectively called "North Oscan." Best attested is **Paelignian**, known from about two dozen inscriptions in the Abruzzi region of eastern Italy; it is thought to be a form of Oscan by some, though its exact affiliation is not clear. Not far away were a cluster of other quite similar languages: **Marsian**, in which we have close to a dozen inscriptions; **Marrucinian** (a half-dozen); **Vestinian** (two inscriptions); and one inscription each in **Aequian**, **Sabine**, and **Volscian**, the latter spoken to the south of Latin.

Furthermore, we have about a half-dozen inscriptions from Campania from the sixth and fifth centuries BC in a language called "Pre-Samnite," which may have been a form of South Picene.

13.79. Farther to the south, in Sicily, were spoken a number of languages whose exact position within Italic (if they are indeed Italic languages) is unclear; very little survives in them, and much of what does survive is incomprehensible. Noteworthy is **Sicel** (or **Sikel** or **Siculian**), spoken over much of Sicily; an important form preserved on a cup is the imperative *pibe* 'drink!', which is exactly cognate with Sanskrit *piba* 'drink' and has not undergone the assimilation of the first consonant seen in Latin *bibe*.

For Further Reading

Several books treat the history of Greek and Latin together; two were mentioned in the previous chapter (Meillet and Vendryes 1968 and Sihler 1995). The most accessible book in English on the history of Latin has been Palmer 1954, which has an appendix of archaic texts; to this may now be added Baldi 1999. The standard comparative grammar of Latin is Leumann 1977, exhaustive but largely eschewing laryngeals; more up-to-date, though smaller, is Meiser 1998. A detailed investigation of the outcomes of the laryngeals in Latin is Schrijver 1991. The finest etymological dictionary of Latin is Ernout and Meillet 1979, which contains a wealth of cultural information alongside the purely linguistic. Most of the important Archaic Latin texts are collected, with linguistic commentary, in Ernout 1947; see also, more recently, Wachter 1987 and Vine 1993.

A vast resource on the antiquities of ancient Italy is the multi-volume *Popoli e civiltà dell'Italia antica* (Rome: Biblioteca di Storia Patria, 1974–92), of which vol. 6 contains descriptions of all the languages. The most up-to-date collection of all the Sabellic inscriptions is Rix 2002, which contains the raw texts with references to secondary literature but no commentary. The standard collection of all the Italic dialect inscriptions known by the early 1950s is Vetter 1953, which has interlinear Latin translations and some commentary; it was supplemented by Poccetti 1979. The Faliscan corpus is in Giacomelli 1963. For Oscan and Umbrian, still very useful because of its thorough grammatical descriptions is Buck 1928, which contains most of the important inscriptions plus translations and a glossary. Of the many editions of the Umbrian Iguvine Tables, Poultney 1959 is the most accessible, and has an English translation. Meiser 1986 is a useful recent phonological history of Umbrian. A dictionary of Sabellic is the new Untermann 1999. The South Picene corpus is handsomely edited in Marinetti 1985.

For Review

Know the meaning or significance of the following:

Etruscan "Italo-Celtic"	rhotacism Festus	Vulgar Latin Appendix Probi	Sabellic Iguvine Tables
----------------------------	---------------------	--------------------------------	----------------------------

Exercises

- Using the sound and morphological changes introduced in this chapter, determine the Latin outcomes of the following PIE forms. In *a* and *e*, the final *-i* disappeared prehistorically.

a	* <i>h₂eǵonti</i> 'they drive'	g	* <i>duk₂tos</i> 'led'
b	* <i>ph₂tres</i> 'of the father'	h	* <i>str₃h₃tos</i> 'laid, strewn'
c	* <i>h₃sih₁mos</i> 'may we be'	l	* <i>g^wi₃h₃uos</i> 'alive'
d	* <i>tr₃h₂ns-pr₂teh₂tiō</i> 'act of carrying across'	j	* <i>ek₂uom</i> 'horse' (accus.)
e	* <i>deuketi</i> 'he leads'	k	* <i>ṇter</i> 'between'
f	* <i>loukejoh₂</i> 'I make shine'		

- The outcomes of the PIE voiced aspirated stops in Latin are fairly complex. Using the data below, answer the questions that follow.

* <i>bhāgos</i> > <i>fāgus</i> 'beech'	* <i>h₁leudheros</i> > <i>līber</i> 'free'
* <i>bhrātēr</i> > <i>frāter</i> 'brother'	* <i>h₁rudhē-</i> > <i>rubā-</i> (<i>facere</i>) 'to redden'
* <i>bheremos</i> > <i>ferimus</i> 'we bear'	* <i>ǵhiems</i> > <i>hiems</i> 'winter'
* <i>h₃erbhos</i> > <i>orbus</i> 'orphan'	* <i>ghostis</i> > <i>hostis</i> 'stranger, enemy'
* <i>albhos</i> > <i>albus</i> 'white'	* <i>ghabhē-</i> > <i>habē-</i> 'have'
* <i>dhu₂h₂mos</i> > <i>fūmus</i> 'smoke'	* <i>ghu-n-d-</i> > <i>fund-</i> 'pour'
* <i>dhightos</i> > <i>fictus</i> 'fashioned by hand'	* <i>ghyer-</i> > <i>fera</i> 'wild animal'
* <i>medhijos</i> > <i>medius</i> 'middle'	* <i>ueǵheti</i> > <i>uehit</i> 'conveys'
* <i>uidheueh₂</i> > <i>uidua</i> 'widow'	* <i>dhingh-</i> > <i>fin-</i> 'make with the hands'
* <i>oudhes-</i> > <i>ūber-</i> 'udder'	* <i>g^whormos</i> > <i>formus</i> 'warm'
* <i>joudhejoh₂</i> > Archaic <i>ioubeo</i>	* <i>sning^wheti</i> > <i>ninguit</i> 'it snows'
'I command'	* <i>snig^whm</i> > <i>niuem</i> 'snow'

- a Treating **gh* and **ǵh* as the same thing, determine the word-initial outcomes of each aspirated stop.
- b Again treating **gh* and **ǵh* as the same thing, determine the word-internal outcomes of each aspirated stop.
- c It was mentioned in §2.15 that Eng. *law* comes from the root 'to lay down', a root that has also been forwarded as the source of the Latin word for 'law', *lex* (stem *leg-*). The root in question is **legh-*. Why is this etymology of the Latin word problematic?
- 3 In the Introduction (§1.1) and elsewhere in this book it has been mentioned that Lat. *deus* 'god' is not related to its Greek synonym *theós*. Explain why this is so.
- 4 Explain the historical relationships between the words in each of the following pairs or triplets. For example, if the pair were *dīcit* '(s)he says' : *dictus* 'said' (past participle), the answer would be that the *i/i* alternation probably continues PIE ablaut, namely full grade (present tense, thematic) and zero-grade (**to-* verbal adjective), which means *i* is from **e/i* (§13.35). Not all the answers involve matters of PIE date.
- a *ūrīt* '(s)he burns', *ustus* '(having been) burned'
- b *precēs* 'prayers' (nomin. pl.), *proculus* 'suitor'
- c *tegīt* '(s)he covers', *toga* 'toga' (< "covering"), *tēgula* 'roofing tile'
- d *fīdīt* '(s)he trusts', *foedus* 'treaty', *fīdēs* 'faith, trust'
- e *iungīt* '(s)he joins, yokes', *coniugem* 'spouse' (accus. sg.; *con-* 'with')
- f *fūner-is* 'of a funeral (gen. sg.)', *fūnes-tus* 'funerary'
- 5 In the Latin *suovitaurlia* prayer (§13.53), you met the forms *-que* and *ē-uen-īre*, where the *-qu-* and *-u-* (which stands for the glide *w*) go back to IE **kʷ* and **gʷ*, respectively. Now consider three other Latin words:
- bōs* 'cow'
- coquīna* 'kitchen' (earlier **quoquīna*)
- popīna* 'hash-house'
- The latter two forms go back to an Italic root **kʷekʷ-* or **kʷokʷ-* meaning 'cook'. Which of these three forms (*bōs*, **quoquīna*, *popīna*) show(s) the genuine Latin outcome(s) of the labiovelars in question? What explanation can you offer for the forms that do not?
- 6 a Analyze the excerpt from the *suovitaurlia* prayer (§13.53) in light of the strophic poetic style discussed in §§2.42–44. Note especially pairs or triplets of words that echo each other phonetically or syntactically.
- b Cato's text has *bonam* in the last line of the excerpt; as the Notes indicate, this is from earlier **duenam*. If we restore **duenam* here, how would that fit in with the phonetic figures that you just discussed?
- 7 Latin has pairs such as the following, in which the first word is a derivative with a **no-* suffix and the second word is a diminutive of the first: *tignum* 'plank' : *tigillum* 'small plank'; *pugnus* 'fist' : *pugillus* 'handful' (< "little fist"). Latin also has

pairs where the first member has undergone sound change, such as *scamnum* 'stool' : *scabillum* 'low stool'.

- What would *scamnum* have originally been, and what sound change occurred to give the attested form?
- Another pair where the non-diminutive has undergone sound change is *pānis* 'bread' : *pastillum* 'medicinal tablet' (< *'small loaf'). What would *pānis* have been originally, and what sound change(s) occurred to give the attested form?
- Latin also has the pair *pīlum* '(large) pestle' (with an *-i*-suffix instead of an *-n*-suffix) : *pistillum* '(small) pestle'. What would *pīlum* have been originally, and what sound change(s) occurred to give the attested form? (Ignore §13.7 for the purposes of this problem.)
- Comparable to the above is the pair *āla* 'wing' : *axilla* 'little wing'. What would *āla* have been originally, and what sound change(s) happened to give the attested form?
- How does the sound change you came up with for d explain the difference between the prefix in the verbs *ex-portāre* 'to carry out' and *ē-lūcere* 'to shine out'?
- What additional sound change is needed to explain the compound verb *ē-numerāre* 'to count out'?
- An inscription from the third century BC has the form *losna* (from slightly earlier **lousna*) 'moon', which became Classical Latin *lūna*. This is from the same root seen in *lūc-ēre* 'to make shine' < **louk-ēre*. What does this inscriptional form tell us about the order of events making up the sound change in f?

8 Observe the following pairs of Latin forms:

ag-ere 'to do' : *āc-tus* '(having been) done'
fra-n-g-ere 'to break' : *frāc-tus* '(having been) broken'
leg-ere 'to choose' : *lēc-tus* '(having been) chosen'

The lengthened vowels in the past participles are unexpected. Contrast *fac-tus* '(having been) made' from *fac-ere* 'to make' and *iac-tus* 'thrown' from *iac-ere* 'to throw'. It has been claimed that in Latin, a vowel before a voiced stop became lengthened when that stop was devoiced before a voiceless consonant ("Lachmann's Law"). What rule in PIE phonology might pose difficulties for this claim?

- How do we know that *duenoi* in the third line of the Duenos inscription (§13.52) had a long diphthong *-ōi* rather than *-oi*, as per the notes?
- Using the information in §13.13, indicate into which of the four Latin conjugations the following PIE verb forms would have fallen:

<ol style="list-style-type: none"> *<i>h₁rudh-eh₁-je-</i> 'be red' *<i>doḱ-éje-</i> 'show' 	<ol style="list-style-type: none"> *<i>pek^w-e-ti</i> 'cooks' *...-eh₂-je-
--	---
- Based on §13.13 and your knowledge of PIE and Latin sound changes, into which of the four conjugations would the following PIE athematic verbs have fallen?

<ol style="list-style-type: none"> *<i>bhleh₁-ti</i> 'weeps' 	<ol style="list-style-type: none"> *<i>bheh₂-ti</i> 'speaks' 	<ol style="list-style-type: none"> *<i>neh₁-ti</i> 'sew'
--	--	--

PIE Vocabulary V: Body Functions and States

- **g^w/h₃yo-* 'alive': Ved. *jīvā-*, Gk. *bíos* 'life', Lat. *uiuus*, Eng. **quick**
- **h₂enh₁-* 'breathe': Ved. *ániti* 'breathes', Gk. *ánemos* 'wind', Lat. *anima* 'breath'
- **syep-* 'sleep': Ved. *svápnas*, Gk. *húpnos*, Lat. *somnus*
- **syeld-* 'SWEAT': Ved. *svédate* 'sweats', Lat. *sūdor*
- **h₁ed-* 'EAT': Hitt. *ēdmi* 'I eat', Ved. *ádmi* 'I eat', Gk. *édomai* 'I will eat', Lat. *edō* 'I eat'
- **peh₃-* 'drink': Hitt. *pāši* 'swallows', Ved. *pāti* 'drinks', Lat. *pōtus* 'a drink'
- **ġenh₁-* 'give birth': Ved. *jánati* 'gives birth', Gk. *génesis* 'birth, beginning', Lat. *genus* 'race, kind'
- **yeid-* 'see': Ved. *ávidat* 'found', Gk. *é(w)idon* 'I saw', Lat. *uideō* 'I see', OCS *viděti* 'to see'
- **k₁leu-* 'hear': Ved. *śṛṇóti* 'hears', Gk. *klŭthi* 'hear!', OIr. *ro-cluinethar* 'hears', Lith. *klausýti* 'to hear'
- **men-* 'think': Ved. *manyáte* 'thinks', Gk. *maínomai* 'I go mad', Lat. *mēns* (*ment-*) 'mind', Eng. **MIND**
- **u₁emh₁-* 'vomit': Gk. (*w*)*émein* 'to vomit', Lat. *uomere* 'to vomit'
- **perd-* 'FART': Ved. *párdate* 'farts', Gk. *pérdetai*
- **mer-* 'die': Gk. *ámbrōtos* 'immortal', Lat. *mors* (*mort-*) 'death', Eng. **MURDER**

14 Celtic

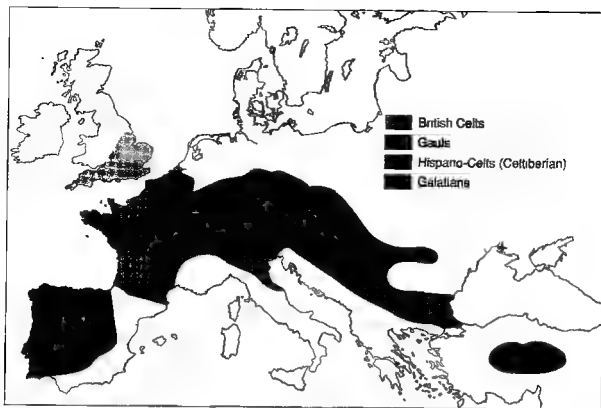
Introduction

14.1. The Celtic languages hold a special place in the early history of Indo-European linguistics because they presented the first real challenge to the nascent science. The demonstration that Irish and its relatives are related to the likes of Greek, Latin, and Sanskrit was a genuine triumph; for while it is obvious that Greek, Latin, and Sanskrit are related to each other, it is not at all obvious that they have anything to do with Irish or Welsh – languages that, on the surface at least, are bafflingly different. We will discuss how the puzzle was solved below when we talk about Insular Celtic (§§14.21ff.).

14.2. The Celtic languages that have survived in unbroken tradition until the present day are confined to a small corner of northwestern Europe – Irish Gaelic in Ireland, Scottish Gaelic in Scotland, Welsh in Wales, and Breton in Brittany (northwest France); the total number of their speakers does not exceed one million. Such meager numbers give little indication of the erstwhile glory of this branch of Indo-European. For hundreds of years before the expansion of late republican Rome in the first century BC, Celtic tribes dominated much of Europe. Archaeologically, it appears that the prehistoric Celts are to be identified with the later stages of the Hallstatt culture (c. 1200–500 BC), located in what is now southern Germany, Austria, and Bohemia (western Czech Republic). By the end of this period, Celtic tribes had spread outward in almost all directions, first westward into France, Belgium, Spain, and the British Isles, and then, by about 400 BC, southward into northern Italy and southeast into the Balkans and beyond, with one group (the Galatians) eventually winding up in Asia Minor (see further below).

After Julius Caesar's conquest of Gaul (ancient France) by 50 BC and the emperor Claudius's subjugation of Britain roughly a century later, most of this Celtic-speaking territory was assimilated to the Roman world. Latin became the dominant language; Gaulish and the other Continental Celtic languages eventually died out. The other branch of Celtic, *Insular Celtic*, to which all the modern Celtic languages belong, continued to flourish in the British Isles, especially in Ireland, whose separation from Britain by the Irish Channel insulated it somewhat from the Romans and, later, from the Anglo-Saxons. Ireland in fact is the home of the first vernacular literature written in medieval Europe, that is, literature that was not written in an official language of the Church (Latin in the West).

But eventually the expansion of English cultural hegemony reduced the number of Celtic speakers in Ireland and elsewhere in the British Isles. The situation is no



Map 14.1 The Celts in the late first millennium BC

better in France, where Breton, spoken by descendants of British Celts, suffers under the comparable cultural dominance of French.

Celtic and other branches of Indo-European

14.3. Celtic shares several features with Italic, leading some scholars to claim that the two branches formed an "Italo-Celtic" subgroup of Indo-European. But the validity of this claim is in doubt, even after decades of controversy. See §13.5 for a more detailed discussion.

From PIE to Celtic

Phonology

STOPS

14.4. Celtic is a centum branch, having merged the palatal velars with the ordinary velars. A defining change was the loss of **p* in most positions, as in OIr. *athair* 'father' < **ph₂tēr*. At an early date, the voiced labiovelar **g^w* became *b* (e.g. **g^wen-* 'woman' > OIr. and W. *ben*); interestingly, the other labiovelars remained intact as labiovelars until much later. Following the change of **g^w* to *b*, the voiced aspirates lost their aspiration, as in OIr. *biru* 'I carry' < **bher-oh₂*, Middle Ir. *daig* 'fire' < **dheg^wh-i-*, and Middle W. *gell* 'yellow' < **g^hel-*.

A double-dental sequence (§3.36) became *-ss-* in Celtic (pronounced as a single *s* by the historical period): OIr. *-fess* 'known' < **uid-to-*, W. *guŷs* 'summons, writ' < **uid-tu-*.

LARYNGEALS

14.5. Laryngeals were lost except when vocalized, in which case they became **a*, as in Italic and some other branches (e.g. **ph₂tēr* > OIr. *athair*, cp. Lat. *pater*, Eng. *father*).

RESONANTS

14.6. The nonsyllabic resonants stayed unchanged except for final **-m*, which became *-n* in Gaulish and Insular Celtic. The syllabic liquids are a difficult domain of Celtic phonology because of their multiple outcomes. Sometimes **ɾ* **l* became **ri* **li*: OIr. *críde* 'heart' < **k₂rd-ijo-* (cp. Gk. *kardiā*); Gaul. *litano-* and W. *llydan* 'wide' < **p₂ltano-* (cp. Gk. *platús* 'broad'). But the outcomes **ar* **al* are also found: Gaul. *Arto-*, W. *arth*, OIr. *art* 'bear' (whence the name Art) < **h₂rtko-*; OIr. *tart* 'thirst' < **t₂rsto-* (cp. Ved. *tṛṣṭá-* 'dry', Eng. *thirst*).

The syllabic nasals are more straightforward: **m* and **n* became **am* and **an*: Gaul. *ambi-* 'around', W. *am* < **h₂m₂bhi*; OIr. and W. *an-* 'not, un-' < **n-*. In Irish, as recent research has shown, these were sometimes raised to *em/en* or *im/in*, especially before voiced stops, as in *imb* 'around'.

14.7. The long syllabic resonants **r̥* **l̥* **m̥* **n̥* (from sequences of syllabic resonant plus laryngeal; see §3.15) typically turn into the relevant resonant followed by *ā*: Gaulish city name (Latinized) (*Medio-)*lānum 'middle of the plain, Milan' < **p₂lh₂no-* (cp. Lat. *plānum* 'plain'); OIr. *grán* 'grain' (the acute accent indicates length), W. *grawn* 'grain' < **gr̥h₂no-* (cp. Lat. *grānum* 'grain', Eng. *corn*); OIr. *gnáth* 'known, customary' < **gn̥h₂to-* (cp. Gk. *gnōtós* 'known').

VOWELS

14.8. The IE vowel system remained largely unchanged. The main early shift was made by **ō*, which became **ū* in final syllables, as in the 1st sing. ending in Gaul. *delgu* 'I hold' and OIr. *biru* 'I carry'. It became **ā* elsewhere: **mōros* 'great' > Celtic **māros*, as in the Gaulish personal name *Sego-māros* 'great in strength', Archaic OIr. *mār* 'great', and W. *maur* 'great'; cp. Gk. *enkhest-mōros* 'great with the ash spear'.

Morphology and syntax

VERBS

14.9. Most of the verbal formations of PIE survive into Celtic in one form or another, including athematic verbs, aorists, perfects, thematic subjunctives, and middles. The optative, though, has disappeared, as have participles (though a few old participles are preserved vestigially as nouns, such as OIr. *carae* 'friend' from the present participle **karant-* 'loving'). Dual verb endings were also lost (but not dual noun endings), although there is disputed evidence of a dual in Gaulish. The Celtic imperfections are of unknown origin and must be innovations. Like Italic, Celtic has a derivational suffix in *-ā-* from PIE **-eh₂-* (§5.34) for forming denominative verbs, as well as a subjunctive morpheme in *-ā-* of uncertain origin (cp. §13.19).

NOUNS

14.10. The new finds from Celtiberian have shown that Celtic inherited the PIE case-system intact in the singular. Like Italic, Celtic has an o-stem genitive singular in *-i (e.g. Primitive [inscriptional] Ir. *maq(q)i* 'of the son').

SYNTAX

14.11. Insular Celtic is unusual among the older IE branches in that verbs are generally clause-initial. (The possible source of this order was discussed in §8.19.) An interesting and archaic feature of Gaulish and Insular Celtic word order was the position of the relative pronoun. In these languages, an indeclinable relative *yo (from the relative pronominal stem *iō-, §7.10) was placed after the first element in a relative clause. This is attested directly in such forms as Gaul. *dugiionti-io* 'who serve', and (somewhat obscured after sound changes) Old W. *issid* (Mod. W. *sydd*) 'who is' < *esti-yo. In Celtiberian, however, declined forms of the relative pronoun are still found, such as the dative *iomui* 'for whom', and its placement seems to have been freer.

Continental Celtic

14.12. The Celtic languages spoken in continental Europe until the first few centuries AD, all of them extinct, are referred to as Continental Celtic and were spoken by people called *Keltoi* by the Greeks and *Galli* or *Galatae* by the Romans. Chief among the Continental Celtic languages is Gaulish. The Gauls were a powerful Celtic people consisting of hundreds of tribes spread throughout much of Europe; as noted above, their territory stretched from Gaul (ancient France) through Switzerland and northern Italy into Hungary, with one major group settling in the third century BC in Galatia in central Anatolia. (The *Gal-* of Galatia is related to *Gaul*; its inhabitants are the Galatians addressed in an epistle of Paul in the New Testament.) Many of the Gaulish tribes were aggressive marauders, as exemplified by the *Galatians*, who wreaked havoc on many principalities in Asia Minor, and famously by the band of Gauls that penetrated deep into Italy in the early fourth century BC and sacked Rome in 390 or thereabouts. But the tides of history eventually turned against them; they were subdued by Julius Caesar by 50 BC and gradually assimilated into Roman culture. Gaulish was still spoken in isolated pockets probably until around AD 500, or even later in Asia Minor.

Gaulish inscriptions are written in both the Greek and Italic scripts. The number of known inscriptions has grown dramatically since the mid-twentieth century. In addition to the inscriptional remains, many Gaulish proper names are preserved in Greek and Roman historical and geographical writings.

14.13. Inscriptions in a Celtic language called *Lepontic* (also called Cisalpine Celtic) have been found in northern Italy; recent scholarship regards it as a dialect of Gaulish. The earliest Lepontic inscriptions date to the sixth century BC and represent our oldest preserved Celtic. Most of the Lepontic inscriptions are funerary and contain little more than personal names.

14.14. Last of the Continental Celtic languages is Celtiberian (or Hispano-Celtic), spoken by Celtic peoples who migrated into northeast Spain in the mid-first millennium BC. Most of the nearly 120 extant Celtiberian inscriptions are very brief;

the language has been seriously studied only since 1970, when an extensive bronze inscription was found in the village of Botorrita (ancient Contrebia Belaisca), 20 km south of Zaragoza in northeast Spain. An even longer second inscription was found at the same site in 1992, but disappointingly consists almost entirely of personal names. About three other inscriptions of some length are known, two from south of Zaragoza and one that recently came to light in New York City and was published in 1993. The Celtiberian inscriptions date to the second and first centuries BC.

14.15. The script of most of the Celtiberian inscriptions was borrowed from a neighboring non-Indo-European people, the Iberians, who lived along the eastern Spanish coast and acquired an alphabet from either the Greeks or the Phoenicians or both. They modified it in certain ways so that it became partly a syllabary: vowels and resonants were indicated by single letter signs, while stop consonants were only indicated by signs that represented the combination of a stop plus a following vowel. No distinction was made between voiced and voiceless stops, so a sign like *ta* stood for both *ta* and *da*.

14.16. A very poorly attested IE language called Lusitanian is documented in Portugal. Although some say it is a relative of Celtiberian, this is very uncertain. Lusitanian will be discussed in chapter 20. On the southernmost tip of the Iberian peninsula was spoken another poorly known language called Tartessian, which has some Celtic linguistic material in its personal names.

Continental Celtic grammar

14.17. We cannot compile anything resembling a complete grammar of either Gaulish or Celtiberian. It does not appear that the two languages were particularly close. Of the two, Celtiberian was the more conservative; as noted above (§14.10), it preserved PIE nominal inflection practically intact. Even the ablative and instrumental, which had merged in Hittite and Greek a millennium earlier, are still preserved in Celtiberian separately as *-uz* and *-u* in *o*-stem nouns (< late PIE **-ōd* and **-ō*). The dual is still a living category in nouns. Celtiberian also preserved the labiovelars as such (as in the conjunction *-cue* 'and' < **k^we*), while Gaulish changed them to labials (as in *pissiumí* 'I see' < **k^wi-*). Since Brittonic also changed the labiovelars to labials (§14.54), some specialists view Gaulish as a Brittonic language, but this claim has not won general acceptance.

Gaulish text sample

14.18. An excerpt from the inscription found in Chamalières in Clermont-Ferrand (southeast France), written on a lead tablet in the early first century AD and deposited in a spring. In ancient Mediterranean cultures, this was standard practice for prayers, curses, and other texts addressed to underworld deities. This tablet was written apparently on behalf of some men seeking beneficial action from the gods. There are no word-divisions in the original, and not all the forms are understood, so no translation is given.

[...] toncnaman
toncsíontío meíon ponc sesit bue-
tid ollon reguc cambion exsops
pissiumí [...]

14.18a. Notes. *toncnaman*: a noun in the accus. sing. forming an etymological figure with the following verb; perhaps 'oath'. *toncsíontio*: a 3rd pl. verb *toncsíont(i)* 'they will swear' (?), containing the future morpheme **-sje-* (§5.36), with the relative pronominal clitic *-fo* 'who' or 'which' attached (PIE **jo-*). Perhaps '... the oath which they will swear ...'. *ponc*: 'when' or the like, from **kwom-k(e)* (accus. sing. of the relative pronoun plus the deictic particle **ke*, §7.28); cp. Lat. *tunc* 'then' < **tom-k(e)*. *buetid*: 3rd sing. *buet* 'will be' plus clitic subject pronoun *-id* 'it'. *ollon*: 'all, whole'. *reguc cambion*: probably 'I straighten the crooked'; *regu* is 1st sing. from the root **h₃reg-*, seen also in Lat. *regō* 'I direct', and *cambion* is from the Celtic root **kamb-* 'crooked', also referring to back-and-forth motion and exchange. It is ultimately the source of Eng. *change* via Late Latin *cambiāre* 'to exchange', a borrowing from Celtic. The reason for the *-c* at the end of *reguc* is disputed. *exsops*: 'blind', literally '(having) the eyes (*-ops*) out (*exs-*)'. It has been suggested by Calvert Watkins that the French word for 'blind', *aveugle*, from Vulgar Latin **ab-oculus*, is ultimately a loan-translation of this Gaulish word. *pissiumi*: 'I will see', 3rd sing. future, from pre-Celtic **k^wid-sjō(-mi)*. The root is the source of OIr. *(ad-)cī* 'sees'. The last four words are thus '(And) I (shall) straighten the crooked; blind, I shall see'.

Celtiberian text sample

14.19. The first three lines of the first Botorrita inscription. The transliteration makes use of the most recent discoveries, using *b* for the signs formerly transliterated with *p*, and using *s* and *z* instead of *ś* and *z* (*z* represents either [z] or [ʒ]). A very tentative translation is given, mostly following Wolfgang Meid, *Celtiberian Inscriptions* (1994).

tiricantam bercunetacam tocoitoscue sarniciocue sua combalcez nelitom
necue to uertaunei litom necue taunei litom necue masnai tizaunei litom soz aucu
arestalo tamai uta oscuez stena uerzoniti silabur sleitom conscilitom cabizeti

Concerning the region (?) pertaining to the **berguneta* of Tocoit- and Sarnicios, it has been thus decreed as non-permitted:

It is neither permitted to put (things) upon, nor is it permitted to do (work?), nor is it permitted to cause damage by destruction.

And whoever contravenes these things, he shall take cut-up ... silver ...

14.19a. Notes. *tiricantam*: uncertain; translated by Meid 'region', but it has been recently suggested to be the number 'thirty' in the sense of a group of thirty people (a council?) or other societal division (cp. the division of the Roman people by Romulus into thirty divisions). Cp. OIr. *trichat* < Celt. **tri-kant-* 'thirty' < PIE **tri-dkmt-*. *bercunetacam*: phonetically probably *bergunetacam* and apparently containing the root for 'hill', *berg-* (PIE **bherǵh-*, cp. German *Berg* 'mountain'); maybe 'hilly'. *tocoitoscue sarniciocue*: 'both of T. and of S.'; *-cue* ... *-cue* are like Lat. *-que* ... *-que* 'both ... and', from PIE **k^we*. *sua combalcez*: contains the notion 'it has been thus decreed' or the like, but grammatically unclear; could also be read *compalcez*, *complacez*. *Sua* may be comparable to Goth. *sua*, Eng. *so*; *combalcez* may be a 3rd sing. past tense verb with ending from **-t*. The exact phrase *tiricantam* ... *sua combal[ce]z* recurs on another Celtiberian inscription. *nelitom*: 'not allowed', from the same root as Eng. *let*. *necue* ... *necue*: 'neither ... nor', cp. Lat. *neque* ... *neque*. *to uertaunei*: perhaps literally 'the putting-upon', if phonetically to *uerdaunei*, with *uer-* 'over, above' (cp. the Gaulish name *Ver-cingeto-rix* 'far-stopping king') and *daunei* an infinitive meaning 'to do, place, put' from **dheh₁* 'place, put', similar in form to Gk. (Cypriot) *dowenai* 'to give' (from a different root). *litom*: 'allowed'. *masnai*: dative, perhaps from a noun **mad-sna-* related to OIr. *maidid* 'breaks'. *tizaunei*: another infinitive. The rest of the line and the beginning of the next are unclear. The reading *arestalo* is uncertain (*arestaso* is also possible). *uta*: 'and'. *oscuez*:

'whoever', formally like Gk. *bós-tis* < **ios-kʷis*. *uerzoniti*: 'over-seeks, contravenes', another verb with the prefix *uer-*. *silabur*: 'silver', a cultural loanword also found in Germanic and Balto-Slavic (cp. Goth. *silubr*, Russ. *serebro*). *cabizeti*: 'he shall take'.

Insular Celtic

14.20. Insular Celtic consists of two subbranches, Goidelic and Brittonic. Goidelic contains Old Irish and its descendants: Irish Gaelic, Scots Gaelic, and Manx. Brittonic (or Brythonic) contains Welsh, Breton, Cornish, and perhaps the language of the Picts (see §14.53). (Manx and Cornish both died out in recent times, though there have been attempts to revive them.) Goidelic and Brittonic are often called "Q-Celtic" and "P-Celtic" because of their respective treatment of the PIE labiovelars: Goidelic turned the labiovelars into velars, while Brittonic turned them into labials.

While Continental Celtic languages look basically like other old Indo-European languages, with Insular Celtic the story is quite different. Irish and Welsh, to put it bluntly, look bizarre when compared to languages like Greek and Latin. This is mostly due to a massive set of sound changes that occurred in rather quick succession over the course of a few centuries, and to which we now turn.

Phonology

INSULAR CELTIC CONSONANT MUTATIONS

14.21. The single most striking feature of the Insular Celtic languages from the early medieval period onward is the system of initial consonant mutations: word-initial consonants change depending on what word precedes, or depending simply on grammatical roles. To take Modern Welsh as an example, the word *car* 'car' can appear as *car*, *gar*, *nghar*, or *char*: *eu car nhw* 'their car', *y gar* 'the car', *fy nghar i* 'my car', *ei char hi* 'her car'. Similarly, in Old Irish, the number *sé* 'six' causes no change in a following consonant (*sé guth* 'six voices'), whereas *cóic* 'five' causes a following stop consonant to become a fricative (written *cóic gotha* 'five voices' but pronounced [ʲoθə] with a voiced velar fricative), while *secht* 'seven' causes the consonant to become nasalized (*secht ngotha* 'seven voices'). The system seems utterly capricious, the contexts in which mutation occurs completely arbitrary.

14.22. But behind many capricious linguistic systems lie less capricious origins. We owe the unraveling of this particular mystery to Franz Bopp, one of the founders of IE linguistics in Germany in the early nineteenth century. Until his work, there was no consensus that the Celtic languages were even Indo-European. Bopp's fundamental breakthrough was to notice systematic correspondences between the Celtic mutations and certain facts in Greek, Latin, and Sanskrit. Specifically, an Irish word caused a following consonant to become a fricative when the Sanskrit cognate form ended in a vowel, and an Irish word caused nasalization when its cognate ended in a nasal; in all other cases, no mutation happened. For instance, consider the mutations induced (or not induced) by the Old Irish forms below in light of the final sounds of their Sanskrit and Latin cognates:

Skt. *asya* 'his'
eṣām 'their'
asyāḥ 'her'

OIr. *a guth* [ʲoθ] 'his voice'
a nguth 'their voice'
a guth 'her voice' (no mutation)

Lat. <i>quīnque</i> 'five'	ÓIr. <i>cóic gotha</i> [ʋoθə] 'five voices'
<i>septem</i> 'seven'	<i>secht ngotha</i> 'seven voices'
<i>sex</i> 'six'	<i>sé gotha</i> 'six voices' (no mutation)

The mutations were thus revealed to preserve vestiges of final syllables in prehistoric Celtic: the nasalization in a phrase like *a nguth* is a relic of the nasal that once ended the possessive *a* 'their'. (In effect, the mutations are remnants of old sandhi rules.) The overall system is the same in the other Insular Celtic languages, though differing in detail. The mutation that causes a following consonant to become a fricative is called **lenition**, and the mutation that adds a nasal before a following sound is called **nasalization**. The mutations also happened word-internally.

INSULAR CELTIC VOWEL CHANGES

14.23. The "old Indo-European" look of Insular Celtic was also dramatically altered by a series of changes to the vowels, in particular the loss of most final syllables (apocope), the loss of internal vowels (syncope), and umlaut (called in Celtic studies vowel "affection" and "infection," and sometimes "vowel harmony"). Some of these changes happened in the common Insular Celtic period, but others were later parallel and independent developments (arising no doubt from some favorable initial conditions already present in common Insular Celtic).

14.24. The various umlauts were the first to take place. Their details, as well as those of the initial consonant mutations, are discussed individually for Irish and Welsh below. Following the umlauts, both Goidelic and Brittonic underwent apocope and syncope. Syncope in Irish hit every other syllable after the initial (stressed) syllable, e.g. pre-ÓIr. **cossamil* 'similar' > ÓIr. *cosmil*; **ancossamili* 'dissimilar' (pl.) > ÓIr. *écsamli*. In Brittonic, syncope hit unstressed syllables directly before stressed syllables, but was a bit more sensitive to word-structure, as in the British king's name *Cunobelinus* becoming **Cun-belinos* (whence Welsh *Cynfelyn* and English *Cymbeline*), rather than expected **Cuno-blinos* (which would have become **Cyneflyn*).

Morphology

VERBS

14.25. Much inherited IE verbal morphology is preserved intact in Insular Celtic, generally more faithfully in Goidelic than in Brittonic. Both Irish and Welsh continue the *s*-aorist as two types of preterite, an *s*-preterite (e.g. ÓIr. *mórais* 'he magnified', Middle W. *kereis* 'he loved') and a *t*-preterite (e.g. ÓIr. *birt* 'he carried', Middle W. *cant* 'he sang'), the latter being a special development of the *s*-preterite in resonant-final roots. The PIE perfect is continued by another preterite, the reduplicated preterite (although Brittonic barely has any examples): Middle W. *ciglef*, ÓIr. (ro-) *chúala* 'I heard', both ultimately from PIE **ke-klou-*. (In the Irish form, the diphthong *úa* is ultimately from the sequence **ek-* in **keklou-*.) Both Irish and Welsh have a so-called *s*-subjunctive, from the PIE *s*-aorist subjunctive (e.g. ÓIr. indicative *guidiu* 'I pray', *s*-subjunctive *gess*; Middle W. indicative *car* 'loves', subj. *carho* [with *-h-* < **s-*]). Insular Celtic has a 3rd person passive used as an impersonal, while the **to*-verbal adjective is continued as a passive preterite (e.g. ÓIr. *-crith* 'was bought' < **k^wrih₂-to-*; Middle W. *caffat* 'was had'). The IE future (or desiderative) formations

were mostly lost in Brittonic, but survive as futures in Irish (§14.41). A morphological innovation in the Insular Celtic verb is the use of the preverb **ro-* to mark perfective aspect (OIr. *ro-*, W. *ry-*; from IE **pro-*).

The hallmark of the verb in the older stages of Insular Celtic is the distinction between the so-called absolute and conjunct verb-forms: most verb-forms existed in two variants, one used clause-initially and the other when certain elements preceded. This feature is far more characteristic of Irish than Brittonic, and will be discussed in the Irish section.

INFIXED PRONOUNS

14.26. Object pronouns in the Insular Celtic languages are often inserted between the preverb and verb, and are called infixed pronouns. Thus in Old Irish, to say 'does not strike' one said *ní-ben*, while 'does not strike me' was *ní-m ben* (with the initial *b-* pronounced as the fricative [β]) and 'does not strike him' was *ní-mben* (where *mb* was pronounced [mb]). Similarly, from *ad-cí* 'sees' one could form *atút-chí* 'sees you' and *at-chí* 'sees it'. As forms like *ní-mben* show, the presence of an infixed pronoun is sometimes betrayed only by the mutation of the initial consonant of the verbal root (after the preverb).

CONJUGATED PREPOSITIONS

14.27. Characteristic of Insular Celtic is the fusion of prepositions with personal pronominal objects, resulting in prepositions that are "conjugated" in the three persons, singular and plural, like verbs. An example from Middle Welsh, the conjugation of the preposition *ar* 'on', is given below; the forms *arnam* and *arnunt* have taken over endings from the verbal conjugations:

	<i>sing.</i>	<i>pl.</i>
1	<i>arnaf</i> 'on me'	<i>arnam, arnan</i> 'on us'
2	<i>arnat</i> 'on you'	<i>arnawch</i> 'on you'
3	<i>arnaw</i> 'on him'	<i>arnadui, arnunt</i> 'on them'
	<i>arnei, erni</i> 'on her'	

Goidelic: Old Irish and Its Descendants

History of Irish

14.28. Proto-Goidelic, the prehistoric ancestor of Irish, was spoken in Ireland at least by the beginning of the Christian era, if not earlier. The earliest preserved Irish is the language of the 300 or so stone inscriptions written in the Ogam (or Ogham) alphabet, which consists of strokes and notches chiseled along and across a central line, usually the edge of a stone (unfortunately the part that gets weathered and damaged most easily). The origin of the Ogam script is not known for certain; the inscriptions hail mostly from southern Ireland and date from about the fourth to the seventh centuries AD, with the bulk of them coming from the fifth and sixth. The stage of Irish at this time is called Primitive Irish (or Ogam Irish). Its earliest specimens do not yet show the effects of apocope, syncope, or vowel affection (e.g.

maq(q)i 'of the son', OIr. *maicc*; *Lugudeccas* 'of Lugudecca', OIr. *Luigdech*; *velitas* 'of a bard', OIr. *filed*). While this allows us to date certain sound changes that happened during or later than the Primitive Irish period, the value of the inscriptions is offset by their extremely inconsistent spelling and their brevity, as they are all short burial inscriptions consisting almost solely of proper names.

The conversion of Ireland to Christianity in the fifth century resulted in the introduction of the Roman alphabet. Irish clerics learned the alphabet from monks in western Britain, who spoke early Welsh or its immediate ancestor. The fifth and sixth centuries saw not only significant alterations to the cultural landscape of the Emerald Isle, but also radical changes in the Irish language. Broadly, it was during this time that Irish changed from looking roughly like Gaulish or Latin to looking like Irish. The Old Irish period begins with the earliest datable literature, written at the end of the sixth century or the start of the seventh but preserved only in much later manuscripts. The early part of this period, into the first quarter of the seventh century, is called **Archaic Old Irish**, followed by **Classical Old Irish**, which lasted until the mid-900s.

14.29. Several manuscripts preserved on the Continent containing glosses on the Scriptures and other Latin texts represent the most important source of information on Old Irish. The major ones are the Würzburg, Milan, and Turin glosses. These manuscripts were brought to the Continent by Irish missionaries and other religious figures in the eighth and ninth centuries and are our only contemporary documentation of the Classical language. The Würzburg glosses are written very accurately, and their spelling is considered to be the Classical norm. Because the glosses were not understood on the Continent, these manuscripts were not used in later centuries and survived in good shape. By contrast, manuscripts of the same age do not survive from Ireland itself since they became worn out from continuous use. Copying and recopying often introduced corruptions and modernized spellings. These medieval (and even early modern) manuscripts are the ones that contain all preserved Old Irish literature, such as the saga of the *Táin Bó Cuailgne* (*The Cattle-Raid of Cooley*) and other tales, heroic poetry, and legal, historical, and grammatical texts.

14.30. Lasting from the tenth to the thirteenth centuries was Middle Irish, which saw changes to the morphological system that were as far-reaching as the phonological changes of the Primitive Irish period. By the end of the thirteenth century the language was effectively as it is today. The **Modern Irish** period begins with the codification of a normative form of the language by bards and other literary elite in the thirteenth century. The bards' literary guild collapsed in the early 1600s, after which texts in different regional varieties appear – the dialects of Munster, Connacht, and Ulster that are still with us.

14.31. It was also during the seventeenth century that Ireland received an English-speaking ruling class; by the next century the status of Irish had deteriorated and it became primarily a language of the rural poor. A tragic blow was the devastation wreaked by the Potato Famine in 1845–49, which ravaged especially this population; over a million died outright, and another million and a half emigrated to America or elsewhere. Today, no more than 70,000 people comprise the Irish-speaking communities called the *Gaeltacht*. While a larger number have learned Irish as a second language, its future is cloudy.

Irish is often called "Gaelic," though this is not strictly speaking correct. "Gaelic" is properly used in combination, as in "Irish Gaelic" or "Scottish Gaelic"; it is a general term that is more or less synonymous with Goidelic. (It derives, in fact, from the Scottish Gaelic descendant of Old Irish *Góidél*.)

Phonological developments of Old Irish

CONSONANTS

14.32. The labiovelars of Goidelic were changed to ordinary velars during the early history of Irish. Ogam Irish has a separate letter for k^w that is transliterated Q, as in the frequent word *maq(q)i* 'of the son' < **mak^(w)k^wi* (Classical OIr. *maicc*), and a special name is preserved for another letter that must have represented g^w before it shifted to g .

Lenition and nasalization are the main changes to have affected consonants during the historical period. Both changes were conditioned by the preceding sound; the conditioning sound could be in the same word as the affected consonant, or separated from it by a word-boundary. But not just any word-boundary was "invisible" for the purposes of lenition and nasalization – only those in particular syntactic or prosodic groups. Thus lenition or nasalization could apply across the boundary between noun and a following adjective, and after proclitics of various kinds – the definite article, possessive pronouns, certain conjunctions, and infixed pronouns; but they could not apply across two different phrases.

14.33. **Lenition.** Lenition turned originally intervocalic stops into fricatives, as in *cath* 'battle' < **katus* and in *chinn* 'of the head' < **sindi k^wenni*. In the case of voiced stops, the difference between the lenited and unlenited versions is not indicated in the orthography, hence *bó* 'cow' begins with *b*, but in *a bó* 'his cow' it begins with a bilabial fricative [β] (similar to Eng. *v*). The fricatives *s* and *f* when lenited became *h* (written *š*) and zero (written *f*), and the nasals and liquids when lenited were lenis (less strongly articulated) versions of the corresponding unlenited nasals and liquids. These facts are all still true of modern Irish dialects.

14.34. **Nasalization.** The other consonant mutation, nasalization, affected only stops and vowels. It occurred when the final sound of the preceding word was a nasal consonant in prehistoric Irish. Nasalization had three effects. It changed voiced stops to nasal stops, for example *dán* 'gift' but in *ndán* 'the gift' (accus. sing.; *nd* = [n]) < pre-Irish **sindan dānan*. It changed voiceless stops to voiced stops (not represented in the spelling), for example *túath* 'people' but in *túath* 'the people' (accus. sing.), pronounced in *dúath*, ultimately from earlier **sindān toutān*. Finally, it prefixed an *n* to a vowel, for example *ech* 'horse' but in *n-ech* 'the horse' (accus. sing.) < earlier **sindan ek^wan*. The only other mutation to be noted is that a word-initial vowel is prefixed with *h* following a word that does not lenite or nasalize, as in *úa hAirt* 'grandson of Art, O'Hart'.

14.35. **Palatalization.** A front vowel (*i* or *e*) following a consonant palatalized that consonant, that is, moved its point of articulation closer to the palate. In the early stages of this change, only *i*'s in certain syllables had this effect; as time went by, palatalization spread to other contexts. The end result was that consonants, aside from being lenited, unlenited, and nasalized, now came in two other guises, palatalized and non-palatalized (or "slender" and "broad," as Celticists refer to them). Palatalization is indicated in the orthography by writing an *i* before the relevant consonant (but this is not done consistently): for example, *-beir* 'carries', with palatalized (and lenited, since postvocalic) *r* that used to be followed by a front vowel (**bhereti*; cp. the absolute form *berid*). The *b* is also palatalized because it occurs before the vowel *e*, but this is not specially shown in the script.

VOWELS

14.36. The main Irish changes to the Common Celtic vowel system are as follows. An *o in an unstressed syllable became *a*, as in the Ogam consonant-stem genitive singular *Lugudeccas* 'of Luigdech', with *-as* < *-os. These *a*'s mostly disappeared but not before causing *a*-affection (see below). The *u*-diphthongs **eu* and **ou* were monophthongized to *ó* (as in Archaic OIr. *tóth* 'tribe' from **toutā*), later diphthongized to *úa* (hence classical *túath*). The diphthong **ei* became *é* or (after the archaic period) *ía* depending on whether the vowel that originally stood in the following syllable was front or back, respectively: thus *tégi* or *-téig* 'you (sing.) go' < *(s)teigh-es(i) but *tiagu* 'I go' < *(s)teigh-ō.

14.37. **Vowel affection.** The vowels were also subject to various umlauting processes called vowel affection. The details of vowel affection are highly complicated and will not concern us here in detail. There were two basic processes in Irish: *a*-infection and *i*-infection. In *a*-infection, a high vowel (*i* or *u*) was lowered to the corresponding mid vowel (*e* or *o*) before a syllable containing an *a*; this *a* usually came from an earlier **o* (see previous section). Thus nomin. sing. **uir*-os 'man' > **uir*-as > **uer*-as (ultimately OIr. *fer*; contrast the nominative pl. *fir* < **uir*-ī); **klut*-om 'fame' > **klut*-an > **klot*-an (ultimately OIr. *cloth*). *i*-infection produced the opposite result; an *i* caused an *e* or *o* in a preceding syllable to raise to *i* or *u*: *mil* 'honey' < **melit*; *suide* 'seat' < **sodijom*.

The subsequent changes of apocope and syncope were discussed above, §14.24.

Morphology of Old Irish

VERBS

14.38. The Old Irish verb is a complex subject; only a few topics pertaining to it can be treated here. There were five indicative tenses (present, imperfect, future, secondary future [also called past future or conditional], and preterite), two subjunctive tenses (present and past), and a present imperative. Active and middle endings were distinguished in the primary (non-past) tenses, and there were distinct passive endings in the third person. Special relative endings are found in the 3rd singular and plural and in the 1st plural; these were used when the verb was in a relative clause, and are historically derived from a combination of verb plus relative pronoun (§14.11 above). Instead of infinitives, each verb had a verbal noun formed from the verbal root with a suffix, somewhat like the different infinitives of Vedic (§10.44).

14.39. **Present stems.** There were two classes of verbs, distinguished according to their present stems: "strong" (all formed directly from verbal roots) and "weak." The strong verbs are old and continue present-stem types of PIE date: simple thematic presents, such as *berid* 'carries' (< **bhereti*); (thematized) nasal-infixes presents (e.g. *dingid* 'oppresses'); and *-na*-presents comparable to the Sanskrit ninth class (§10.42; e.g. *crenaid* 'buys', cp. Ved. *krīṇāti*). Weak verbs are secondarily derived presents such as causatives and denominatives, such as *móraid* 'increases' from *mór* 'great'. All original athematic verbs, except for some forms of the verb 'be', have become thematic.

14.40. **Other tense-stems.** There were four other stems for the other verb forms: subjunctive, future, preterite active, and preterite passive. In the strong verbs, as

with the primary verbs of other IE languages like Sanskrit and Greek, the other tenses are formed directly from the root.

In contrast to most of the other older PIE languages, the present subjunctive is not formed from the present stem. The subjunctive stem is formed in two ways: with a morpheme **-ā-* that is identical to the *ā*-subjunctive of Italic (§13.19; e.g. pres. indicative *renaid* 'sells', pres. subjunctive *-ria*, cp. Lat. *attingit* 'touches', Archaic Lat. subjunctive *attigat* without nasal infix), or with an *-s-* (§14.25 above).

14.41. The future stem is formed in a variety of ways. Weak verbs take the *f*-future, whose origin has been hotly debated for over a century (e.g. pres. *léicid* 'leaves', fut. *léicfid* 'will leave'). Strong verbs generally have a reduplicated future, either with a suffix *-s-* (e.g. from *guidid* 'prays' is formed *gigis* 'will pray' < **g^whi-g^whedh-s-*) or without (e.g. pres. *canid* 'sings', fut. *cechnaid* < **ke-kan-a-*; *do-gní* 'knows', fut. *do-gén* < **gi-gn-a-*; this last type is called the long-*e* future). Both of the reduplicated future types go back either directly or indirectly to the PIE reduplicated desiderative (§5.37).

14.42. The active preterite is formed with an *-s* or with reduplication (there are some other types as well which we shall not go into). All weak verbs have an *s*-preterite (§14.25). The reduplicated preterite, a continuation of the IE perfect (§14.25), is formed by many strong verbs: *canid* 'sings', perfect *cechan-*. Finally, there is a passive preterite stem, which is a continuation of the **-to*-verbal adjective (§14.25).

14.43. Absolute and conjunct verb forms. Verbs in Old Irish (and occasionally in Old Welsh) come in two different guises depending on their position in the sentence. One form appears when the verb is clause-initial, such as OIr. *berid* 'he carries'; this is the absolute form. A generally shorter form, sometimes lacking the last syllable of the absolute form, appears when the verb is preceded by certain conjunctions or a preverb, as in *ní beir* 'he does not carry', *do-beir* 'he gives'. This is called the conjunct form.

Compound verbs themselves – preverb and verb both – behave in the same way as uncompound ones: a full form (such as *do-beir* above) occurs clause-initially, while a shorter form occurs when certain conjunctions or another preverb precedes (as in *ní tabar* 'he does not give'). The terminology is different: the full form *do-beir* is called the proterotonic form, while *-tabar* is called the deuterotonic. The deuterotonic is thus sort of a conjunct of the conjunct.

14.44. The origin of the differences between absolute and conjunct forms is not altogether agreed upon, but it probably boils down to something like the following. One must first assume that in pre-Irish, clauses contained an unstressed clitic sentence particle in second position, i.e., following the first element in the clause. (This assumption is not terribly problematic since this is in one way or another true of Greek and Hittite, and frequently true in Sanskrit and elsewhere.) If the first element in the clause was an uncompound verb, the particle attached to the verb's final syllable; when the language was hit by apocope, which got rid of most final syllables, the particle was treated as a final syllable and protected the rest of the verb in front of it. Thus a **beret(i)* + Particle became the absolute form *berid*. With compound verbs, one needs to remember that preverbs were independent adverbs originally (§8.9); therefore if a compound verb began a clause, the preverb would have been the first element in the clause and the sentence particle would have attached to it rather than to the following verb. The final syllable of the verb was therefore not protected and fell victim to apocope. Thus a **to* + Particle *beret(i)* became *do-beir*, with *-beir* the conjunct form.

If a compound verb was itself preceded by another element, such as the negative **nē*, similar principles applied. The clause would have started **nē* + Particle *to-beret(i)*. Here we must further assume that the compound **to-beret(i)* acted like a single word in such an environment, with stress on the first syllable like all other words in pre-Irish. The difference in stress between **to* + Particle *béreti* and **nē* + Particle *tó-beret(i)* is what mattered for the subsequent phonological developments: unstressed internal syllables were subject to syncope (§14.24) and other vowel weakenings. Thus **tó-beret(i)* became deuterotonic *-tabar*.

14.45. As these examples hint at, the effects of syncope and apocope in Irish are nowhere as apparent or as devastating as in verbal conjugations, especially those of verbs compounded with one or more preverbs. The effects are particularly widespread partly due to the fact that Old Irish had an especial fondness for piling preverbs together, giving syncope and apocope no shortage of syllabic gallows-fodder. Three preverbs in a row are quite common, e.g. *du-airmgerat* 'they promise' from **do-are-in-gar-* (verbal root *gar-* 'call', literally 'call to before in' or the like; the literal translations of these compounds usually make no sense), and four or even five preverbs in a row are found as well, e.g. *fo-timmdiriat* 'I am sufficient' from **fo-to-imm-di-ret-* (verbal root **ret-* 'run'; literally 'run under to around from') and the verbal noun *contherchomracc* 'assembly' from **com-to-er-com-ro-icc-* (verbal root **icc-* 'come, reach'). Because only the first or second preverb was stressed (depending on its position in the clause), all syllables after that were unstressed and got compressed by lenition, syncope, apocope, and vowel weakening in unstressed syllables. The verbal root itself wound up leading an extremely perilous existence at or near the end of such strings, and in some cases all but entirely disappeared.

14.46. The amount of allomorphy (that is, variation in the form of a given morpheme) that these changes created was incredible, and it is worth digressing to give some examples. The Irish root *féd-* 'say' appears in such varied guises as *fét*, *id*, and *d*, as in the following forms of the compound verb **at-féd-* 'relates': 3rd sing. present *ad-fét* 'relates', perfective present *ad-cuid* (**ad-com-féd-*), and 2nd pl. conjunct perfective present *-éicdid*. Sometimes the root is reduced to a single sound, particularly in the conjunct 3rd person singular *s*-subjunctive. Thus the 3rd sing. conjunct *s*-subjunctive of the verb *as-boind* (**oss-bond-*) 'refuses' is simply *-op* (pronounced *-ob*, the regular outcome of **-óss-bod-s-t!*); and from **ret-* 'run' (present *rethid* 'runs') we have the compound *do-fúarat* 'remains over' (**di-fó-uss-ret-*), whose conjunct 3rd sing. *s*-subjunctive is *-diúair* (< **-di-fo-uss-ret-s-t!*, with only the *-r* remaining of the original root, subjunctive suffix, and personal ending (and very little left of the preceding preverbs)).

The effects of these sound changes caused absolute and conjunct forms even of relatively simple verbs to be vastly different. Thus the pre-Irish compound **di-slond-it* 'denies' becomes OIr. *do-shuindi* (absolute) but *-díltai* (conjunct); both developed by perfectly regular sound change. The absolute form comes from **di-slóndit* (accented on the verb), whereas *díltai* comes from the same form accented on the preverb (**-dí-slondit*). The destressing of *-slondit* made all the difference: **di-slondit* first became **di-sldi* (syncope and loss of final **-t!*), then **dibhti* (weakening of **s* to **h* which devoiced the following consonant cluster), and finally *dílti* (loss of *h* with compensatory lengthening of preceding vowel), written *-díltai*.

14.47. All of this allomorphy seems not to have posed any particular problems for Old Irish children learning the language; but as time went by the allomorphy

was gradually reduced by regularization – one or another allomorph was generalized throughout the paradigm by leveling or analogy. The distinction between absolute and conjunct was lost in Middle Irish. Analogy was already a significant player in the paradigms of scores of Old Irish verbs; part of the challenge of Irish historical grammar is figuring out which forms are analogical creations and which are not.

14.48. Below are given the present, imperfect, present *ā*-subjunctive, and *t*-preterite of the verb *berid* ‘carries’. The conjunct forms are illustrated with the compound verb *as-beir* ‘says’, whose deuterotonic forms are given in the third column (-*epur*, as would appear for instance in *ní epur* ‘I do not say’):

Present				Imperfect		
	Absolute	Conjunct	Deuterotonic	Absolute	Conjunct	Deuterotonic
sg. 1	<i>biru</i>	<i>as-biur</i>	<i>-epur</i>	<i>no berinn</i>	<i>as-berinn</i>	<i>-eprinn</i>
2	<i>biri</i>	<i>as-bir</i>	<i>-epir</i>	<i>no berthar</i>	<i>as-berthar</i>	<i>-eperthar</i>
3	<i>berid</i>	<i>as-beir</i>	<i>-epir</i>	<i>no bered</i>	<i>as-bered</i>	<i>-epred</i>
pl. 1	<i>bermai</i>	<i>as-beram</i>	<i>-eprem</i>	<i>no bermis</i>	<i>as-bermis</i>	<i>-epermis</i>
2	<i>berthe</i>	<i>as-berid</i>	<i>-eprid</i>	<i>no berthar</i>	<i>as-berthar</i>	<i>-eperthar</i>
3	<i>berait</i>	<i>as-berat</i>	<i>-epret</i>	<i>no bertis</i>	<i>as-berdis</i>	<i>-epertis</i>
<i>ā</i> -Subjunctive				<i>t</i> -Preterite		
	Absolute	Conjunct	Deuterotonic	Absolute	Conjunct	Deuterotonic
sg. 1	<i>bera</i>	<i>as-ber</i>	<i>-eper</i>		<i>as-biurt</i> ²	<i>-epurt</i> ²
2	<i>berae</i>	<i>as-berae</i>	<i>-epre</i>		<i>as-birt</i>	<i>-epirt</i>
3	<i>beraid</i>	<i>as-bera</i>	<i>-eprea</i>	<i>birt</i>	<i>as-bert</i>	<i>-epert</i>
pl. 1	<i>bermai</i>	<i>as-beram</i>	<i>-eprem</i>		<i>as-bertammar</i>	<i>-epertmar</i> ²
2	<i>berthae</i>	<i>as-beraid</i>	<i>-eprid</i>		<i>as-bertid</i>	<i>-epertaid</i> ²
3	<i>berait</i>	<i>as-berat</i>	<i>-epret</i>	<i>bertatar</i> ¹	<i>as-berditar</i>	<i>-epertatar</i> ²

¹ Late. ² Not actually attested.

NOUNS

14.49. Irish reduced the number of cases in the noun to five – nominative, vocative, accusative, genitive, and dative. Below are the paradigms of the *o*-stem *fer* ‘man’, the *ā*-stem *túath* ‘tribe’, and the consonant-stem *carae* ‘friend’ in the singular and plural (the superscript ^L means the form lenites a following word, and ^N means it nasalizes):

	Sing.	Pl.	Sing.	Pl.	Sing.	Pl.
nom.	<i>fer</i>	<i>fir</i> ^L	<i>túath</i> ^L	<i>túatha</i>	<i>carae</i>	<i>carait</i>
voc.	<i>á fir</i> ^L	<i>á firu</i>	<i>á thúath</i> ^L	<i>á thúatha</i>	<i>á charae</i>	<i>á chairtea</i>
acc.	<i>fer</i> ^N	<i>firu</i>	<i>túait</i> ^N	<i>túatha</i>	<i>carait</i> ^N	<i>cairtea</i>
gen.	<i>fir</i> ^L	<i>fer</i> ^N	<i>túaithe</i>	<i>túath</i> ^N	<i>carat</i>	<i>carat</i> ^N
dat.	<i>fiur</i> ^L	<i>feraib</i>	<i>túait</i> ^L	<i>túathaib</i>	<i>carat</i> ^L	<i>cairtib</i>

Forms that nasalize a following word once ended in **-m*, and forms that lenite once ended in a vowel. The vocative is always preceded by the particle *á* 'O'. Unlike any other Indo-European language, the vocative plural is distinguished from the nominative plural in the *o*-stems (and, by analogy, in the consonant stems), which happened for the following reason. The inherited nominative/vocative plural **-ōs* was replaced (as in some other branches) by the pronominal nomin. pl. **-oi* (§6.53), but interestingly only in the function of nominative plural; the old ending **-os* lived on in its other function as vocative plural. The replacement of a form in its main function but not in one of its secondary functions is a common process, and goes by the name of Kurylowicz's Fourth Law of Analogy, after the Polish linguist Jerzy Kurylowicz.

Old Irish text sample

14.50. §§17–21 of the Old Irish wisdom text *Audacht Morainn* (Testament of Morann), after the edition of Fergus Kelly (Dublin, 1976). The language is Archaic Old Irish. The text was written around 700, and consists of advice to a young king by a legendary judge named Morann. (The genre is called *Speculum Principum* or 'mirror of princes'.) The ethical notion of the ruler's truth is of common IE patrimony, and is found in Vedic and later India, ancient Iran, and Greece; in all these societies the verbal expression of this concept is believed to ensure prosperity and protect from harm.

§17 Is tre fír flathemon ad- manna mármeso márfedo -mlasetar.

§18 Is tre fír flathemon ad- mlechti márbóis -moinigter.

§19 Is tre fír flathemon ro-bbí cech etho ardósil imbeth.

§20 Is tre fír flathemon to- aidble éisc i sruthaib -snáithet.

§21 Is tre fír flathemon clanda caini cain-tussimter.

(17) It is through the justice of the ruler that abundances of great tree-fruit of the great wood are tasted. (18) It is through the justice of the ruler that milk-yields of great cattle are maintained. (19) It is through the justice of the ruler that there is abundance of every high, tall corn. (20) It is through the justice of the ruler that abundance of fish swim in streams. (21) It is through the justice of the ruler that fair children are well begotten.

14.50a. Notes. 17. *Is*: 'is', PIE **h₁esti*, pronounced [ɪʃ]. It introduces a cleft sentence ("It is . . . that . . ."), a favorite Insular Celtic sentence type used for topicalization. In the case of this particular sentiment, though, we may be dealing with a pre-Celtic pattern, since a construction beginning "by the ruler's truth" plus a sentence is structurally equivalent to truth-statements elsewhere in IE. *tre*: 'through', leniting the following sound, from the same root as Lat. *trāns* 'across'. *fír*: lenition of *fír* 'justice', literally 'truth', PIE **ueh₁-ro-* (> Lat. *uerus* 'true', German *wahr* 'true'). *flathemon*: 'ruler', genit. sing. of *fla(i)them*. The genitive shows this to be an *n*-stem, Celtic **wlati-amon-* (**wlati-* > OIr. *flaith* 'rule'), ultimately continuing the PIE suffix **-mon-* seen e.g. in Greek. *hēge-mōn* 'leader'. *ad-mlasetar*: 'are tasted', 3rd pl. pres. passive of *ad-mlasat* (otherwise unattested), with archaic *ml-* that turned into *bl-* in Classical Old Irish. Derived from *mlas* 'a taste'. The separation (tmesis) of preverb from verb is an archaic stylistic feature that harks back to the earlier separability of preverb from verb (§8.9) and recurs in the following lines. *manna*: 'abundances', pl. of *mann*, a borrowing

from Lat. *manna* 'manna'. *mármeso*: compound of *már* 'great' (Classical OIr. *mór*; = W. *maur*) and *meso*, genit. of *mes* 'tree-fruit', a *u*-stem cognate with W. *mes* 'acorns'. Acorns frequently symbolized a just rule. *márfedo*: 'of great wood'; *fedo* is genit. sing. of *fed*, also a *u*-stem; Celtic **widu-* < PIE **uidh-*, the source of Eng. *wood*. Note the alliteration in this and the following lines.

18. *ad-móiniger*: 'are treasured, are maintained', 3rd pl. pres. passive of an otherwise unattested verb *ad-móinigerthar*, from *móin* 'treasure' (< **moi-ni-*, from PIE root **mei-* 'exchange', also in Lat. *mīnus* 'service done in exchange for something'). *mlechti*: 'milk-yields' (Classical *blechti*), pl. of *mlicht*, from an old *tu*-abstract noun **mleg-tu-* from PIE **h₂melg-* 'milk, to milk'. In another text, *disce mblechta* 'dryness of milking' is mentioned as one of the proofs of a false king. *márbóis*: 'of great cattle'; PIE **g^wōu-*.

19. *ro-bbí*: 'there is', a form of doubtful interpretation here and perhaps corrupt. *cech*: 'every', genit. sing. of *cach*, with *i*-infection from old genitive ending **-ī*. The word comes from PIE **k^wā-k^wo-*, also seen in Russ. *kakoj* 'which?' etho: 'of grain', genit. sing. of *ith*. Together with its W. cognate *yd*, this continues PIE **pi-tu-* 'nourishment', seen also in Ved. *pītu-* and Av. *pītu-* 'food'. *ardósil*: 'high (and) tall', a compound of *ard* 'high' (cognate with Gk. *orthós* 'upright') and *ósal* (Archaic spelling of *úasal*) 'tall' (Gaul. *uxello-*, from PIE **up-s-*, cp. Gk. *hupsélós* 'lofty'). Various phrases for 'tall grain' are found in traditional IE literatures. *imbeth*: 'abundance', Classical *imbed*.

20. *to-snáithir*: 'swims'; this and the following *aidble* 'abundance' are not otherwise attested and have been restored from somewhat conflicting manuscript readings. Contains preverb *to-* (Classical *do-*) 'to' plus the Irish descendant of PIE **sneh₂-* 'swim' (also > Lat. *nā-re* 'to swim'). *éisc*: 'of fish', genit. sing. of *íasc*; PIE **peisk-* (also in Lat. *piscis*, Eng. *fish*). The manuscripts actually read *uisce éisc* 'fish in water', but modern editors assume *uisce* 'water' was added originally as a gloss and was later accidentally incorporated into the text by a copyist. *i sruthaib*: 'in streams', dat. pl. of *sruth*, originally a *tu*-abstract **sru-tu-* from **sreu-* 'flow', the source of Eng. *stream*.

21. *clanda*: 'children' (Classical *clanna*), pl. of *cland*, cognate with W. *plant*, both early borrowings from Lat. *planta* 'plant' at a time when *p* was heard (and adopted) by Insular Celtic speakers as *k^w*. *caini*: 'fair, beautiful'. *cain-russimter*: *cain-* is 'fair' again, acting as a prefix attached to *-tussimter*, the 3rd pl. pres. passive deuterotonic form of *do-fuissim* 'I beget'. The form *-tussim-* is from **to-uss-sem-*, with two preverbs; the proterotonic form *-fuissim* has an extra preverb *fo-* inserted before *uss-*. The root **sem-* in IE meant 'draw water out of a well' (Lith. *semiù* 'I draw water').

Scottish Gaelic and Manx

14.51. Beginning around the late fourth century, immigrants from Ireland colonized what is now Scotland, establishing a colony called Dál Riata, named after a town in northeast Ireland. The Scottish Dál Riata became the center of a kingdom in the latter half of the fifth century, becoming more powerful than the Irish Dál Riata; in the first half of the seventh century the connection to Ireland was severed.

The Gaelic inhabitants of Scotland expanded inland, and Scottish territory grew to about its present dimensions by about 1000. Over the ensuing centuries, however, English gradually began to replace Scottish, and the predominantly Gaelic-speaking areas are now confined to the Outer Hebrides and the islands of Skye, Tiree, and Islay. No more than 80,000 people speak Scottish Gaelic today, vanishingly few of them monolinguals.

14.52. The same group of Irish colonists that came to Scotland also settled the Isle of Man, in the Irish Sea between Ireland and England, in the fourth or fifth century.

The island's political and cultural heyday came during the period of the Viking invasion and settlement from the ninth to the thirteenth centuries, when it became the center of a Norse-Gaelic kingdom. Subsequently it was ceded to Scotland and then passed to England in the fourteenth century.

Manx is the name given to the form of Goidelic that developed on the island following the Irish colonization. Written Manx begins with a seventeenth-century translation of the Anglican *Book of Common Prayer*; however, a historical poem known as the *Manannan Ballad* is probably a hundred years older, even though its earliest manuscripts are from the eighteenth century. Most of the published literature has been religious in nature.

From the late eighteenth century on, the use of Manx declined, and the last native speaker died in December 1974. It is still in very limited use as a second language.

Brittonic

14.53. Brittonic (also Brythonic or British Celtic) was the language of the Celtic inhabitants of Britain from before the Roman conquest of the British isles beginning in AD 43. At its greatest extent, to judge by the geographical distribution of place-names and river-names, British Celtic was spoken throughout Britain except in that part of Scotland north of the Firth of Clyde and the Firth of Forth. The earliest Brittonic is in the form of some personal names on coins and in scattered Latin and Greek sources, the first being Ptolemy's *Geography* from about AD 150. No connected texts appear until the eighth century, in Old Welsh (see §14.59 below).

In northeastern Scotland lived a people called the Picts (Latin *Picti* 'painted'), whose language, Pictish, has been the subject of controversy. Recent research indicates that Pictish was probably an early Brittonic language, but very little of it is preserved.

14.54. The defining phonological characteristic of Brittonic compared with Goidelic is the change of the Common Celtic labiovelar **kʷ* to *p*, e.g. W. *pump* 'five' < Celt. **kʷenkʷe*, cp. OIr. *cóic*. It is for this reason that Brittonic is often called "P-Celtic." Also characteristic of Brittonic is the change of initial **u* to *gu-*, as in W. *gwr* (OW *guir*) 'man' < **uiros* and *gwyn* 'white' < **uindos* (contrast OIr. *fer* and *finn*).

Vowel affection

14.55. In British Celtic, vowel affection happened before a syllable containing **ā* and before a syllable containing **i*, **ī*, or **ī̄*. Some examples of "ā-infection" include W. *gwen* 'white' (feminine, < **uinda*; compare masc. *gwyn* < **uindos* above) and *Peith(-wyr)* 'Picts' < **Pixtās*. Some examples of "i-infection" (fronting and/or raising) are W. *meibion* 'sons' < **mabjon* and *bryn* 'hill' < **brunnā*.

Mutations

14.56. The initial mutations are somewhat more complicated in Brittonic than in Goidelic. By way of illustration, the Welsh system is summarized in the table below. In Welsh orthography, *ff* represents [f], *f* represents [v], *dd* [ð], *ll* the voiceless lateral [tʃ], and *rh* an aspirated [hr]:

Normal	Lenited	Nasalized	Spirantized
<i>pen</i> 'head'	<i>ar ben</i> 'on top'	<i>fy mhen i</i> 'my head'	<i>ei phen hi</i> 'her head'
<i>tyr</i> 'house'	<i>dy dyr</i> 'your house'	<i>fy nhyr i</i> 'my house'	<i>ei thyr hi</i> 'her house'
<i>cath</i> 'cat'	<i>y gath</i> 'the cat'	<i>fy nghath i</i> 'my cat'	<i>ei chath hi</i> 'her cat'
<i>ben</i> 'woman'	<i>un fen</i> 'a woman'	<i>fy men i</i> 'my woman'	
<i>dyn</i> 'man'	<i>dau ddyn</i> 'two men'	<i>fy nyn i</i> 'my man'	
<i>glaw</i> 'rain'	<i>cot law</i> 'raincoat'	<i>fy nglaw i</i> 'my rain'	
<i>merch</i> 'girl'	<i>un ferch</i> 'a girl'		
<i>llaw</i> 'hand'	<i>ei law</i> 'his hand'		
<i>rhad</i> 'cheap'	<i>cot rad</i> 'cheap coat'		

As can be seen, lenition in British Celtic is similar to Goidelic in the voiced stops, which became voiced fricatives also. (The lenition of *glaw* to *law* is due to a later development whereby [x], the original lenition product of *g*, disappeared.) However, the voiceless stops became voiced stops (rather than voiceless fricatives as in Goidelic). The effect of the nasal mutation on voiced stops was also identical to Goidelic but its effect on voiceless stops was not: the latter are changed into aspirated nasals. Finally, British Celtic has a third mutation not found in Goidelic, the spirant mutation; this resulted from the geminate consonants **pp*, **tt*, and **kk*, which came about from various consonant clusters (e.g. W. *saith* 'seven' < Celtic **septam*). In Goidelic, these simply became unlenited stops, but in Brittonic they became voiceless fricatives.

Welsh

History

14.57. The early Anglo-Saxon settlers of England during the fourth and fifth centuries dubbed the indigenous Celtic tribes the 'foreigners', or *wealas*, the source of our word *Wales*, and their language as *wælisc*, the source of our word *Welsh*. At this time the British Celtic population was linguistically hardly differentiated. Following several major conflicts between them and the ever-expanding Anglo-Saxons in the sixth and seventh centuries, the Celts retreated, and after the Anglo-Saxons had advanced as far west as the river Severn (which runs roughly parallel to the present-day eastern border of Wales), they were split into two groups. One group lay to the northwest, in what is now Wales and a large area to the north; and one to the southwest, in modern Cornwall, Devon, and nearby areas. Many of the inhabitants of this second area had already migrated to northwest France, where their language would develop into Breton (see §14.62); those that stayed in Cornwall spoke what would become Cornish (§14.69). The Celtic spoken to the north of Wales, which must have resembled early Welsh, gradually died out, although traces of it survive even as far north as Cumbria (near Scotland) in the counting systems used by shepherds. One such system, for example, runs *yau tau tethera methera pimp sethera lethera nothera dothera dick*, which is strikingly similar to the Welsh numbers 1–10, *un dau tri pedwar pump chwech saith wyth naw deg*. (These rural counting systems are distantly related to the familiar *eenie meenie miny mo*.)

14.58. A few scattered short inscriptions in what is often called **Primitive Welsh** are preserved from the mid-sixth to near the end of the eighth century. While no literature is extant from this time in its original form, probably some poems by two sixth-century poets, Taliesin and Aneirin, survive in later (and modernized) twelfth- and thirteenth-century versions, most notably Aneirin's elegiac battle poem *Y Gododdin* (*The Gododdin*).

14.59. The earliest piece of connected text is the so-called *Surexit*-memorandum in Old Welsh, a short Welsh and Latin text concerning a lawsuit, dating probably to the end of the eighth century. Though the Old Welsh period lasted until the beginning of the twelfth century, it has left us with only a few other short texts. By contrast, we have very copious literature in **Middle Welsh**, dated from the mid-twelfth century to the end of the fourteenth. The most famous Middle Welsh prose work is the *Mabinogion*, a collection of romances; much poetry was also written, such as *Y Gododdin* mentioned above, as well as histories, grammars, translations of Latin and French literature, and legal texts (many of which drew on much older traditions). The Middle Welsh period closes with the œuvre of Dafydd ap Gwilym (c. 1325–80), a major figure in medieval European poetry, who is transitional to the modern period; his works, and those of writers through the sixteenth century, are often called Early Modern Welsh.

14.60. With the publication of the first Welsh translation of the Bible in 1588 we enter the true **Modern Welsh** era. This translation set a literary standard for generations and was instrumental in continuing the use of the language, which might well have died out otherwise. For already beginning in the Middle Welsh period the higher strata of society were becoming Anglicized, and Welsh was banned from use as an official administrative language in the mid-1500s. (This was not reversed until 1967.) The use of Welsh has slowly declined ever since; about a half-million people still speak it, with the greatest density being in the northwest Welsh county of Gwynedd.

Old Welsh text sample

14.61. The first three of the twelve anonymous "Juvenus" poems (*englynion*), written on a ninth-century manuscript containing a Latin paraphrase of the Gospels by Juvenus. The interpretation and translation below are adapted from Ifor Williams, "The Juvenus Poems" (*The Beginnings of Welsh Poetry*, Cardiff, 1980, 89–121).

niguorcosam nemheunaur henoid mitelu nit gurmaur
mi am [franc] dam ancalaur.

nicanāniguardam nicusam henoid cet iben med nouel
mi amfranc dam anpatel.

namercit mi nep leguenid henoid is discirr micoueidid
dou nam riceus unguetid.

I shall not keep watch even one hour tonight, my retinue is not very large,
I and my Frank, round our cauldron.

I shall not sing, I shall not laugh, I shall not jest tonight though we drank clear
mead,
I and my Frank, round our bowl.

Let no one ask me for merriment tonight, Mean is my company,
Two lords can talk: one speaks.

14.61a. Notes. 1. *niguorcosam* *nemheunaur*: unclear; the translation above is conjectural. *Niguorcosam* is composed of *ní* 'not' and some verb in the 1st sing. present (ending -am) and containing the preverb *guor*- 'above, over' (= OIr. *for*-); if *nemheunaur* does mean 'even one hour', it would be segmented *nemh eun aur*. *henoid*: 'tonight', Modern W. *heno*, earlier *henoeth*; *noeth* 'night' is from PIE **nokʷt-*. *mítelu*: *mí telu* 'my retinue', ModW *fy nheulu* 'my family'; the nasal mutation of the initial consonant was not indicated in the Old W. original. *gurmaur*: 'very large'; *gur*- is the same as *guor*- in *niguorcosam* above, and *maur* (ModW *mawr*) is cognate with OIr. *már*, *mór*. *franc*: 'Frank', also 'foreigner, mercenary soldier'. *dam*: 'around', from **do-ambi*. *ancalaur*: 'our cauldron'; *an* is 'our', and *calaur* is a loanword from Late Lat. *caldāria*.

2. *nicanā*: *nicanam*, 'I do/shall not sing' (with the common medieval notation *ā* to indicate *am*); from **kan-* 'sing' (as also in Lat. *canō* 'I sing'). *niguardam*: 'I do/shall not laugh', ModW *chwarddaf* 'I laugh'. *nicusam*: 'I shall not jest' is a conjectural translation. *iben*: 'we drank', from **pibe-* (cp. Ved. *pibati* 'drinks', Lat. *bibit*). *med*: 'mead', PIE **medhu-* 'honey, sweet drink' (cp. Gk. *méthui* 'wine', Eng. *mead*).

3. *namercit*: 'let no one ask me', containing *na-*, the negative prohibitive particle, -m-, infixed 1st sing. dative pronoun, and the verb *ercit*, a 3rd sing. imperative from PIE **perk-* 'ask' (also in OIr. *com-arc* 'he asks'). Infix pronouns practically disappeared by Middle Welsh times; already here, *mí* 'me' is added after the verb to reinforce the infixed pronoun. *nep*: 'no one', subject, placed after the verb as usual in Celtic. *leguenid*: 'merriment', ModW *llawenydd*. *micoueidd*: 'my company, retinue'; *coneidd* (ModW *cyweithas* with a different suffix) comes from *kyueith* 'friend'. *dou*: 'two'. This last line is of uncertain meaning. *unguetid*: 'one speaks', probably *un* 'one' and *guetid* 'speaks', appearing in ModW in the compound *dyweddy* (-g- disappeared in composition by lenition in Welsh). The sense of the verse is that the speaker has lost all his retinue except one and is therefore in no laughing mood.

Breton

History

14.62. The encounters between the Anglo-Saxons and Britons (the native British Celtic population) were sometimes peaceful, sometimes not. A series of pitched battles between the two peoples in the sixth and seventh centuries ended in the defeat of the Britons. Even before then, the aggressive inland raids of the Anglo-Saxons struck fear into the Celts as far away as southwest England, and in the period around 450–470 a group of them, mostly from Cornwall and Devon, emigrated across the waves to northwest France, where they settled in the area now called Brittany (French *Bretagne*), named after the land of their origin. Over the course of the next century or so there were additional migrations. Their language soon developed in its own way, becoming what we now call Breton (*brezhoneg* in Breton itself, after *Breizh*, Brittany).

14.63. Breton appears to have had its greatest geographical extent in the ninth century, from about the middle of which we have our first specimens of Old Breton

in the form of glosses. From the eleventh century we have our only continuous Old Breton text, four lines buried in a copy of a Latin charter from 821; presumably the Breton is from the time of the copy and not the original.

14.64. From the tenth to the thirteenth centuries, the area in which Breton was spoken shrank by almost one-half; since then it has shrunk slightly further. We have many literary works in Middle Breton (c. 1100–1659, the year of the publication of the first grammar and dictionary), most of which date to after 1450 and comprise religious texts and plays that are translations of French and Latin originals. Native literature in Breton is a relatively late creation, not appearing until the nineteenth century.

14.65. Breton never became the language of any political or cultural center, and no standardization ever developed; modern Breton is composed of the most diverse dialects of any modern Celtic language. Nowadays only the western half of the area corresponding to the erstwhile Duchy of Brittany is Breton-speaking (called Basse Bretagne, as opposed to the eastern and French-speaking Haute Bretagne), and even here the language is restricted to the countryside. Breton has also borrowed much more heavily from French throughout its history than any of the other British Celtic languages ever have from English, to the extent that two-fifths of the ordinary vocabulary is of French origin, according to some estimates.

The use of Breton was actively suppressed by the French government until 1951, when laws banning its teaching were relaxed. Since then there has been a remarkable increase in interest in the language and in materials printed in it. The current number of Breton speakers is perhaps half a million. Nevertheless, there is still barely any official recognition of the language, and its long-term survival is not certain.

14.66. Breton is similar to Welsh and Cornish in most structural respects. Syntactically, it differs by placing the subject before the verb rather than after it. One morphological feature of particular interest is the complexity of the number contrasts in the noun. Aside from the ordinary distinction of singular vs. plural, some Breton nouns can form two plurals, one of which is the default and the other of which emphasizes variety or diversity (e.g. *park* 'park' ~ *parkoù* 'parks' ~ *parkeier* 'various different parks'), while some others can form doubly marked plurals with specialized meanings (e.g. *bugel* 'child' ~ *bugale* 'children' ~ *bugaleoù* 'groups of children'). The plurals of diminutive nouns have the unusual property of being marked twice for plural, once before and once after the diminutive suffix: *bag* 'boat' and *bagig* 'little boat', plurals *bagoù* and *bagoùigoù*. Breton has a large number of collective plurals (a feature shared with Welsh), like *merien* 'ants' and *blev* 'hair', from which singulative nouns, indicating an individual member of the collective entity, are formed: *merienenn* 'an ant', *blevenn* 'a (strand of) hair'. Breton goes beyond Welsh, however, in being able to form singulatives to ordinary plurals, which then have a different shade of meaning from the ordinary singular: *pesk* 'fish' ~ *pesked* 'fishes' ~ singulative *peskedenn* 'single fish (out of a group or mass of fish)'. This singulative itself can then be pluralized: *peskedennoù* 'single fishes (out of a mass)'. Names for parts of the body can be inflected in the dual, by using a prefix etymologically derived from the number 'two', such as *lagad* 'eye' ~ *daoulagad* 'both eyes'; and, true to form, Breton can pluralize duals: *daoulagadoù* 'pairs of eyes'.

14.67. Verbs in the modern language are conjugated in a present, future, preterite, imperfect, present and past conditional, and imperative. The future is the descendant of the old present subjunctive, and occasionally still has modal usage. All these

categories except for the imperative also have an impersonal form, descended from the middle, such as *karer* 'one loves'.

Middle Breton text sample

14.68. Stanza 185 from the Middle Breton poem *Tremenvan an ytron guerches Maria* (*The Journey of the Blessed Virgin Mary*), as preserved in an edition from 1530 and later re-edited in *Poèmes bretons du moyen âge* by Théodore Hersart, vicomte de la Villemarqué (Paris, 1879). The poem is ultimately derived from the medieval Latin legend *Transitus Beatae Mariae Virginis*.

Me pet Doe, Roen tir, euyt hent,	I pray to God, King of the earth, for the journey
Don miret ni hac hon holl querent.	to defend us and all our friends.
Ha nep a cret en Doe, Roe an sent,	And whoever believes in God, King of saints,
A vezo diouguel e pep hent.	will be secure on every journey.

14.68a. Notes. *me*: 'I'. As in the rest of Insular Celtic, the *m*-forms (originally from the oblique stem of the PIE 1st person sing. pronoun) were generalized; the nominative **egoh₂* disappeared. *pet*: 'pray', a borrowing from Lat. *petere* 'to seek'. *Doe*: 'God'; cp. Old W. *diuuv*. *Roen*: 'King (of) the', a contraction of *roe* 'king' (cp. Cornish *ruif*, W. *rhwyf*) plus the definite article 'n. *tir*: 'earth', ultimately from Lat. *terra*. *euyt hent*: 'for the journey'; *hent* 'road, way' is cognate with Cornish *hins*, W. *hynt*. *don*: 'from us', contracted from *da hon*, and going with the resumptive pronoun *ni* 'us' following the verb. *miret*: 'to ward off, defend'. *ni*: 'us', also spelled *ny*; cp. W. *ni*, OIr. *sní*. *hac*: 'and'. *holl*: 'all', same in Welsh. *querent*: 'friends'; cp. W. *ceraint*, OIr. *carait*, from Celt. **karant*- 'loving, caring' (appearing in Gaulish proper names like *Carantomagus*). *nep*: 'whoever'; cp. OIr. *neb*. *a cret*: 'believes'; *a* is a particle used with verbs, and *cret* is cognate with OIr. *cretid* 'believes' and Lat. *credere* 'to believe'. *sent*: 'saints', pl. of *sant*. *a vezo*: 'will be', pres. subjunctive, lenited from *bezo*. *diouguel*: 'secure'; cp. W. *diogel*, from *di*- 'from, without' + *gogel* 'caution, avoidance'. *pep*: 'every'; cp. Cornish *peb* and OIr. *cach*; for etymology, see note to line 19 of the OIr. text in §14.50a.

Cornish

14.69. As noted above, Cornish is most closely related to Breton. It was spoken by the Celtic inhabitants of southwest England who remained after their neighbors migrated to Brittany in the sixth century. Our first documents in Old Cornish are some glosses from the late ninth or early tenth century, which are followed around the year 1100 by our largest Cornish text from this period, the *Vocabularium Cornicum* or Cornish Vocabulary, a Cornish version of a Latin-Anglo-Saxon glossary compiled a century earlier by the English monk Ælfric at Cerne in Dorset. Although the Cornish of the Vocabulary is usually called Old Cornish, linguistically it is more on a par with Middle Welsh or Middle Breton.

14.70. From the fourteenth to the sixteenth centuries are preserved close to 10,000 lines of Middle Cornish, largely in the form of miracle plays translated from English; these represent the bulk of our entire Cornish corpus. The Late Cornish period lasted until the death of the last native speaker of Cornish, traditionally said to have been

one Dolly Pentreath who died in December 1777. A few other people probably remained who spoke the language, but they were surely dead by the close of the eighteenth century. A dedicated effort has been made recently to revive the language; the result cannot be called authentic, since the paucity of our documentation and the inconsistencies in spelling leave many facts about the pronunciation, grammar, and vocabulary unknown.

14.71. A characteristic phonological development of Cornish is the change of word-final dentals to -z (written -s), as in *dans* 'tooth' (cp. W. *dant*).

Middle Cornish text sample

14.72. An excerpt from the Middle Cornish religious drama *The Life of Meriasek, Bishop and Confessor*, as preserved in a manuscript from 1504 but likely copied from an earlier original. The translation is that of Whitley Stokes (*Beunans Meriasek: The Life of Meriasek, Bishop and Confessor*, London, 1872). The speaker is Nudus (the Naked Man).

3067	Mur yv sur ov galarov ha feynt off heb feladov mensen cafus dyweth tek	Great surely are my sorrows, And faint am I without fail, I would fain have a fair end.
3070	ny vyn mernans ov gueles yma orth ov goheles drefen ov boys anbethek	Death will not see me. It is avoiding me Because of my being foul.

14.72a. Notes (selective). *mur*: 'great', cp. W. *maur*, OIr. *mór*, Gaul. *-māros* in personal names. *yv*: 'is, are'. *sur*: an English loanword, as is *feynt* in the next line. *ov*: 'my'. *ha*: 'and', the same as in Breton. *heb*: 'without', same as in Welsh. *gueles*: 'see', cp. W. *gwely*. *yma*: 'is', same as in Welsh. *goheles*: 'avoid', cp. W. *gochelyd* 'avoid, shun'. Both are compounds of *go-* 'under' (cp. Ir. *fo*; PIE **upo*) and **kel-* 'conceal' (cp. Lat. *cēlāre* 'to conceal'). *boys*: a spelling variant of *bos*, 'to be'; cp. W. *bod*. The -y- in the spelling indicates that the *o* is long. Note the change of the final dental to -s as noted above in §14.71. *anbethek*: 'foul' (originally rendered 'loathsome' by Stokes in his edition).

For Further Reading

Celtic studies is well served by reference works on most topics, many of them in English. Lewis and Pedersen 1937 is the only comprehensive comparative grammar in English and is still eminently useful; it is a shorter revised version of Pedersen 1909–13. At that time, not much Continental Celtic was yet known. A more up-to-date (but technical) treatment of various topics in Celtic historical phonology is McCone 1996, an important monograph.

For Gaulish, indispensable is the handsome and just-completed four-volume critical edition of the Gaulish inscriptions, Duval 1985–2002; a useful selection is the colorful survey in Meid 1992. The longer Celtiberian texts, with illustrations and provisional translations, can be found in Meid 1994; a complete edition, together with exhaustive commentary and a glossary, is Untermann 1997.

For Irish, Thurneysen 1980 is the standard comparative grammar, one of the best grammars of any IE language. A mostly complete etymological dictionary of Old Irish is Vendryes and Lambert 1959–. McManus 1991 is a thorough treatment of the Ogam inscriptions. An excellent book devoted to the early history of British Celtic up to the twelfth century is

Jackson 1953. The standard reference grammar for Middle Welsh is Evans 1964. A complete historical dictionary of Welsh is *Geiriadur Prifysgol Cymru* (Cardiff: Gwasg Prifysgol Cymru, 1950–2002; in Welsh and English), which has etymologies and textual citations much like the *Oxford English Dictionary*. A recent technical book on historical Brittonic phonology is Schrijver 1995. Pictish has received an important new treatment in Forsyth 1997.

For Review

Know the meaning or significance of the following:

Hallstatt culture	"P-Celtic"	lenition	Kuryłowicz's Fourth
Franz Bopp	"Q-Celtic"	nasalization	Law of Analogy
Continental Celtic	infix pronoun	vowel affection	spirant mutation
Insular Celtic	conjugated	syncope	
Goidelic	preposition	absolute	
Brittonic	Ogam	conjunct	

Exercises

- 1 Give the Celtic outcomes of the following PIE sounds or sequences. Some may have more than one answer.

a *bh	c *-dt-	e *ō	g *m̥
b *p	d *r	f *gh	h *n̥H

- 2 Briefly explain the history or significance of the following forms:

a OIr. <i>athair</i>	d Gaul. <i>duglioniō</i>	g OIr. <i>á fīru</i>
b Primitive Ir. <i>maqi</i>	e Celtiberian <i>-uz</i>	h OIr. <i>gigis</i>
c Gaul. <i>delgu</i>	f OIr. <i>cath</i>	i W. <i>pump</i>

- 3 Words belonging to the following morphological categories cause lenition of the initial consonant of a following closely connected word in Old Irish. Assuming these categories continue the same categories in PIE, explain why these forms induce lenition.

- a dative singular
- b feminine nominative singular (ignore original final laryngeal)
- c genitive singular of *o*-stems

- 4 Words belonging to the following morphological categories cause nasalization of the initial consonant of a following closely connected word in Old Irish. Assuming these categories continue the same categories in PIE, explain why these forms induce nasalization.

- a genitive plural
- b neuter nominative singular of *o*-stems
- c accusative singular

- 5 Given the phonological rule discussed in §3.40, would you expect the Old Irish nominative singular *cú* 'dog' to mutate the initial consonant of a following closely connected word, and if so, what mutation would you expect to see?
- 6 Given a form like Gaulish *Rigomagus* 'royal field' (*magus* = 'field'), explain why the *s* in the Old Irish compound *rigsuid* 'royal seat' is lenited.
- 7 a Based on the following data, provide an account of the development of clusters consisting of stop plus resonant in Old Irish.

muinél 'neck' < **moniklo-*
mál 'prince' < Ogam *magl-*
canél 'kindred' < **kenetlo-*
(fo-)álagar 'is laid low' < **ad-logar*
én 'bird' < **etno-*
ár 'slaughter' < **agr-*
ám 'a moving back and forth' < **ag-mo-*

- b The Old Irish verb *do-gní* 'does' has a verbal noun *dénam* 'a doing', while the verb *fo-gní* 'serves' has a verbal noun *fognam* 'service'. Which verbal noun shows the historically expected phonological outcome of the word-internal cluster **-gn-*? Provide an explanation for the other form.
- 8 As noted in §14.25, the Insular Celtic *t*-preterite is a special development of the *s*-preterite in resonant-final roots. Given a form like *tart* 'thirst' < **tarst-* (cp. Eng. *thirst*), and given the fact that in various languages new verbal paradigms can be created by adding personal endings to an original 3rd singular form, provide an explanation for the source of the *t*-preterite in Celtic. Hint: do not worry about the ablaut grade of the root.

PIE Vocabulary VI: Natural Environment

- **h₂ster-* 'STAR': Hitt. *ḫašterza*, Gk. *astér*, Lat. *stella*
 **seh₂uōl*, **sh₂un-* 'SUN': Gk. (Homeric) *hēlios*, Lat. *sōl*
 **mēn-s-*, **mēn-ōt-* 'MOON': Ved. *mās*, Gk. *mēn*, Lat. *mēnsis*
 **dheghom-* 'earth': Hitt. *tēkan*, Gk. *khthōn*, Lat. *humus* 'ground', Toch. A *tkam*
 **yod₂r-* 'WATER': Hitt. *wātar*, Gk. *húdōr*, OCS *voda*
 **peh₂ur-* 'FIRE': Hitt. *paḫḫur*, Gk. *pūr*, Umbr. *pir*, Arm. *howr*
 **mori* 'body of water': Lat. *mare*, OIr. *muir*, OCS *morje*, Eng. *mere*
 **doru* 'wood, tree': Ved. *dāru*, Gk. *dōru* 'spear'
 **stenh₂-* 'THUNDER': Ved. *stániti* 'thunders', Lat. *tonāre* 'to thunder'
 **h₂ueh₂nto-* 'WIND': Ved. *vāta-*, Lat. *ventus*
 **nebh-* 'cloud': Ved. *nábhas-*, Gk. *néphos*, Lat. *nebula*
 **sneigh^h-* 'SNOW': Gk. *niph-*, Lat. *nix* (*niu-*), OCS *sněgŭ*

15 Germanic

Introduction

15.1. The branch of IE to which English belongs is called Germanic (or in older literature, Teutonic). The term derives from the name of an ancient Germanic tribe rendered in Latin as *Germānī*; in spite of repeated efforts it still has no accepted etymology. The older name “Teutonic” is also derived from an ancient Germanic tribal name, that of the *Teutonī* or *Teutonēs*, ‘they of the tribe, the people’, from the same root as the word *Dutch* (the language ‘of the people’). Archaeological and linguistic evidence suggests that speakers of Common Germanic lived in northern Europe in the first half of the first millennium BC, primarily in southern Scandinavia and along the coasts of the North and Baltic seas, in an area stretching from the Netherlands in the west to the Vistula River in the east, in what is now Poland. By the time Germanic peoples entered into history, their territory had stretched considerably farther south: the earliest accounts come from the Romans in the first century BC, with whom they would frequently come into conflict. The Roman historian Tacitus left us an important monograph on Germanic peoples and customs, in which he lists the different tribes and describes their religious practices, children’s games, and various other aspects of their culture.

15.2. Contacts with prehistoric Finnic peoples were quite extensive, as evidenced by many words of early Germanic origin borrowed into the common ancestor of modern Finnic languages such as Finnish and Estonian. (These are not Germanic languages, nor even Indo-European, but belong to a separate family called Uralic.) A famous case is Finnish *kuningas* ‘king’, which, remarkably, is still practically identical to its source 2,000 years ago, Germanic **kuningaz* (> Old Saxon *cuning*, Old Eng. *cynig*). The almost uncanny similarity of the modern Finnish form to its ancient ancestor, so apparently frozen in time, has led researchers to mine such loanwords for evidence about the early stages of Germanic. But the degree to which Finnish has actually been a “linguistic icebox” is often overstated, and many of these loanwords are not as trustworthy in detail as they first appear.

The Germanic tribes were also in contact with Balto-Slavic peoples to the east, again as evidenced by loans, such as Slavic **kūņezī* ‘prince’ (> Russ. *knjaz*) from **kuningaz*, and Slavic **xlēbŭ* ‘bread’ (> Russ. *xleb*) from **hlaibaz* (whence Eng. *loaf*).

15.3. Traditionally, the Germanic family is divided into three branches: the extinct **East Germanic**, containing Gothic and the languages of the Vandals, Burgundians, and some other tribes; **North Germanic**, containing Old Norse and its modern Scandinavian descendants; and **West Germanic**, containing English, German, Dutch,

and their relatives. In addition comes **Runic** or **Runic Norse**, the language of the early runic inscriptions, which may tentatively be classified as North Germanic (see §15.34). The interrelationships among the three branches have been the subject of lengthy controversy. The Germanic-speaking area remained a large dialect continuum for probably all its earlier history and much of its later history too, and it can be difficult to tell if a feature shared by any two Germanic subbranches is due to early common innovation or to diffusion of a later local dialectal innovation. Today, the prevailing view is that North and West Germanic formed a subgroup, often called Northwest Germanic. But for most purposes the traditional threefold division of Germanic is fully satisfactory.

15.4. As is customary in Germanic philology, the voiceless fricative sound of Eng. *thin* will be represented by the letter thorn (þ, capital Þ; italic þ, ð) in Germanic reconstructions and in those languages that used this letter (such as Old English and Gothic). The edh (ð) stands for the voiced *th* of Eng. *this*.

As this textbook is written primarily for an English-speaking audience, the section on English is somewhat longer than the sections on the other Germanic languages.

From PIE to Germanic

Phonology

CONSONANTS

15.5. Germanic is a centum branch, having merged the PIE palatal velars with the plain velars.

15.6. Grimm's Law. The sound changes that most set Germanic apart from the other IE branches are the series of changes beginning with Grimm's Law, continuing with Verner's Law, and ending with the shift from mobile pitch-accent to word-initial stress-accent. Grimm's and Verner's Laws are the most famous sound laws in historical linguistics; we shall treat the former first. It is named after the nineteenth-century scholar Johann Jacob Grimm, a lawyer turned Germanic philologist and dialectologist but nowadays most famous for the collaborative fairy-tale collecting he did with his brother Wilhelm. Grimm is one of three pioneering figures in early nineteenth-century IE studies, the others being Franz Bopp (whom we met in the previous chapter) and the Danish philologist Rasmus Kristian Rask. Rask, in fact, was the one who first figured out the sound change now named after Grimm: in 1814, at the tender age of 17, he wrote a prize-winning essay called *On the Origin of the Old Norse or Icelandic Language*, which is a very good comparative grammar of Germanic, Slavic, Lithuanian, and Greek. In it, he explained umlaut (§15.48) for the first time, and noticed the sound correspondences that form the basis for formulating Grimm's Law. Grimm had independently discovered them and was on the verge of publishing them when he chanced across Rask's work; subsequently very influenced by Rask, he published his own version in 1822 in the second edition of his *Deutsche Grammatik* (*German Grammar*).

Grimm's Law is actually not one sound change, but a series of three separate changes which had the net result of shifting all the inherited PIE stops. The first stage, which we can label Grimm I, rendered the PIE voiceless stops into voiceless fricatives; the second stage (Grimm II) changed the voiced stops into voiceless stops;

and the third stage (Grimm III) changed the voiced aspirates into plain voiced stops. Thus:

Grimm I	Grimm II	Grimm III
*p > f	*b > p	*bh > b (b̥)
*t > þ	*d > t	*dh > d (d̥)
*k > h	*g > k	*gh > g (g̥)
*kʷ > hʷ	*gʷ > kʷ	*gʷh > gʷ (gʷ̥), b (b̥)

Most handbooks actually reconstruct the velar fricatives *χ and *χʷ as the outcomes of *k and *kʷ, but in all the older Germanic languages the sounds are written simply *h* or *hw*. (Early Germanic words rendered by the Romans with *ch*, such as the tribal name *Chatti*, are of uncertain phonetic interpretation.) Though we cannot be sure that this *h* did not sometimes represent a velar fricative χ, only in a few cases in the modern daughter languages is that the resultant sound. We will therefore use *h* and *hw* throughout instead of *χ and *χʷ. The outcomes of Grimm III were perhaps in the first instance voiced fricatives (represented by *ð ð̥ g̥ gʷ̥*), but already by the end of the Common Germanic period these had hardened to stops at least word-initially and after nasals.

There seem to have been two outcomes of PIE *gʷh, the expected *gʷ* (as in Goth. *sigggwan* 'sing' < PIE *sengʷh-), but also *b* (as in Eng. *bid*, OE *biddan* < *gʷheid-je- 'pray', and *bane*, OE *banan* 'slayer' < *gʷhonos) and maybe *w*, if Eng. *warm* and its relatives are from PIE *gʷhor-mo- rather than from the rhyme-form *uor-mo- (in the first case, *warm* would be cognate with Lat. *formus* 'hot' and Gk. *thermós* 'heat'; in the second case it would be cognate with OCS *varŭ* 'heat').

The following boldfaced contrasts between English and Latin or Sanskrit consonants are direct results of Grimm's Law (forms not in the same language as the rest of a column are in brackets):

PIE	Latin	English (Grimm I)	PIE	Latin	English (Grimm II)
*ph ₂ tēr	<i>pater</i>	<i>father</i>	*leb-	<i>labium</i>	<i>lip</i>
*tenu-	<i>tenuis</i>	<i>thin</i>	*pod-	<i>ped-</i>	<i>foot</i>
*k ₁ mtom	<i>centum</i>	<i>hundred</i>	*gel-	<i>gelidus</i>	<i>cold</i>
*kʷod	<i>quod</i>	<i>what</i>	*gʷen-	<i>[gunē¹]</i>	<i>queen</i>

PIE	Sanskrit	English (Grimm III)
*bher-	<i>bhar-</i>	<i>bear</i>
*dheh ₁ -	<i>dhā²</i>	<i>do</i>
*magho-	<i>maghām³</i>	<i>[mag⁴]</i>
*sengʷh-	<i>[omphē⁵]</i>	<i>[sigggwan⁶]</i>

¹ Greek, 'woman'. ² 'Put'. ³ 'Wealth'. ⁴ German, 'am able'. ⁵ Greek, 'voice'. ⁶ Gothic, 'sing'.

15.7. Grimm I did not apply if the consonant was preceded by *s*, as in Eng. *star* < PIE *h₂ster- and Eng. *spew* < PIE *speu-. Additionally, if the PIE form had two voiceless stops in a row, only the first one underwent Grimm's Law, as in OE *eaht* 'eight' < PIE *oktō and Eng. *haft* 'handle, hilt' < PIE *kaptō- 'seized'.

15.8. **Verner's Law.** Numerous troubling exceptions to Grimm's Law were disposed of with the discovery by the Dane Karl Verner of the sound change now known as Verner's Law. This sound change affected only the fricatives *f* *þ* *h* and *h'* that had been produced by Grimm I: if they occurred word-internally, and were not immediately preceded by an accented syllable, they became the voiced fricatives *þ* *ð* *g*. Some examples are given below. On the left-hand side are examples of forms that underwent Verner's Law; on the right, for comparison, are forms that did not undergo it because the conditions were not met:

With Verner			Without Verner	
PIE	after Grimm	after Verner	PIE	after Grimm
* <i>upéri</i>	* <i>uféri</i>	* <i>uþéri</i> 'over'	* <i>ápo</i>	* <i>áfo</i> 'off, away'
* <i>ph₂tér</i>	* <i>fapér</i>	* <i>faðér</i> 'father'	* <i>bhráter-</i>	* <i>bróþer-</i> 'brother'
* <i>dekú-</i>	* <i>tehu-</i>	* <i>teǵú-</i> '-ty (10)'	* <i>dék₂m</i>	* <i>téhm₂</i> 'ten'

15.9. Also affected by Verner's Law was the inherited voiceless fricative **s*, which became **z* if it was in the middle of a word and not preceded by the accented syllable:

With Verner		Without Verner	
* <i>kasón-</i>	* <i>hasón-</i>	* <i>hazón-</i> 'hare'	* <i>préusoh₂</i> * <i>frésō</i> 'I freeze'

(This **z* became *r* in English and the other West Germanic languages, as we will discuss later.) Word-final *-*s* often also became **z*, as in Runic *gastiz* 'guest' < PIE **ghostis*; whether this is also part of Verner's Law or an independent change is debated.

Sometime after Verner's Law had run its course the mobile IE accent became fixed on initial syllables (see §15.15 below), which erased the original conditioning factor for Verner's Law. The insight of Verner was to recover the sound change in spite of the erasure of its conditioning factor.

15.10. **Final consonants.** Final stops were lost in words of more than one syllable, as in the PIE optative **uelih₂t* > OHG *wili* 'he wants', Mod. Eng. *will*. Final *-*m* became -*n*, as exemplified below in §15.13.

15.11. **Consonant clusters.** In Common Germanic, **n* before **h* was lost with compensatory lengthening of the preceding vowel. This is the reason why the past tense *brought* does not have a nasal like its present, *bring*. Both forms go back to the dialectal PIE root **bhrenk-* 'bring', which became **brenh-* by Grimm's Law and, in certain forms, **breng-* by Verner's Law. The latter survives today in the present *bring*, while **brenh-* is the basis of the past-tense stem **branht-*, which underwent the loss of nasal to become **brāht-*, whence Eng. *brought*.

One more change worth mentioning is the development of clusters beginning with **z*. In **zd* and **zg*, Grimm II devoiced the stop, and the **z* became *s* by assimilation: **nizdo-* > Eng. *nest*, **mezg-* 'knit' > Germanic **mēsk-* > Eng. *mesh*. Analogously, the clusters **zdh* and **zgh* became **zd* and **zg*, as in **kuzdho-* 'treasure' > Goth. *huzd*, Eng. *board* (showing the further West Germanic change of **z* to *r* mentioned above).

15.12. **Laryngeals.** Germanic has very few traces of laryngeals. Those that became vocalized in Greek, Latin, or Sanskrit were lost in most positions, such as

word-initially before consonant (**h₁s-énti* 'they are' > German *sind*) and after a syllabic resonant (**ǵn̥h₃-to-* 'known' > Gmc. **kunþa-* > OE *cūþ* 'known', Mod. Eng. [un]couth, with the same outcome in Germanic as **ŋ-* 'not' > Gmc. **un-* > Eng. *un-*). However, there are some examples of vocalized laryngeals, as in the word *father*, whose *-a-* continues the laryngeal of PIE **ph₂ter-*.

15.13. Resonants. The resonants all stayed intact, except that in final position an **m* became *n*, as in Eng. *ten* (vs. Lat. *decem* 'ten').

The syllabic resonants developed a *u* in front of them: e.g. PIE **m̥r̥-tro-* 'killing' > Gmc. **murþra-* > Eng. *murder*; PIE **p̥l̥h₃no-* 'full' > Gmc. **fulla-* > Eng. *full*; **ŋ-* 'not' > Gmc. **un-* > Eng. *un-*; **nm̥(m)-ono-* 'taken' > Gmc. **numana-* > OE *numen* > Eng. *numb* and *num-skull*, literally 'taken (as to the senses).' (The *b* in *numb* is a late addition to the spelling and has no etymological significance.)

VOWELS

15.14. Germanic reduced the distinctions among the short vowels from five to four by merging **a* and **o* to *a*, and there was a similar merger in the long vowels but in the opposite direction: **ā* and **ō* merged to **ō*. Thus PIE **ōktō(u)* became German *acht* (and Eng. *eight*, with various later vowel changes), and PIE **bhrātēr* became Goth. *broþar* (with Gothic *o* spelling *ō*), Eng. *brother*.

Traditional historical grammars of Germanic distinguish between two long *ē*-vowels. One is simply the unchanged continuation of PIE **ē* and is sometimes called **ē₁*. The other, termed **ē₂*, was a different sound having several origins, not all of them clear. The two sounds fell together in Gothic, but have divergent outcomes in North and West Germanic: **ē₂* became *ē* (as in the word for *here*, OE and OHG *hēr*, ON *hér*), whereas **ē₁* was lowered to **ā* (as in German *Tat* and ON *dáð* 'deed' < Gmc. **dēðiz* < PIE **dheh₁-ti-*). In English, interestingly, this sound was raised again (*deed*; see §15.57).

Among the various conditioned changes to happen to the vowels in Common Germanic, one may be singled out here, the tendency of **e* to change to *i*. This happened in most environments in Gothic, and in certain environments in the other languages; examples from English include *mid* (cp. Lat. *medius*), *is* (cp. Lat. *est*), and *bind* (PIE **bhendh-*). Other conditioned changes, like umlaut, happened later and will be taken up below.

15.15. The accent shift. As alluded to above, the PIE mobile accent became fixed on the initial syllable in Germanic. Thus after **ph₂tér-* had become **faðér-*, the accent shifted to give **fáder*. But certain prefixes (adverbial preverbs in origin) remained unstressed, as in the English verbs *becóme*, *forǵét*, *mistáke*, *withstánd*.

Morphology

NOUNS

15.16. Germanic retained the nominative, vocative, genitive, dative, accusative, and instrumental cases, although the use of the instrumental is extremely restricted even in the oldest preserved texts. The dual was lost in nouns, but not in the personal pronouns or in the verb. The athematic noun classes of PIE are reduced in number and no longer productive except for the *s*-stems and *n*-stems. Germanic was in fact especially fond of the latter, particularly for forming animate nouns. The old *n*-stem

declension is the source of such English plurals as *ox-en* and *brethr-en*, and of the so-called "weak" declension of animate German nouns such as *Hase* 'hare', oblique and plural *Hasen*. To judge by the fact that English *water* (with *r*) stands alongside ON *vatn* (with *n*), Proto-Germanic inherited the IE word for 'water' as an *r/n*-stem still (§6.31); the same is true of the word for 'fire', Eng. *fire* but Goth. *fon*.

15.17. The grammatical gender distinctions among masculine, feminine, and neuter are preserved. By the regular sound changes affecting vowels (§15.14 above), the PIE *ā*-stem feminines became Germanic *o*-stems, while the thematic masculines and neuters with stem vowel **-o-* became Germanic *a*-stems. Sample comparative paradigms of masculine *a*-stem and feminine *ō*-stem nouns from Gothic, Old High German, Old English, and Old Norse are given below. Representing the *ō*-stems is the noun for 'gift' (and Old Norse *skor* 'edge'); representing the masculine *a*-stems is the noun for 'day'.

<i>a</i> -stem (masculine)					<i>ō</i> -stem (feminine)			
	Goth.	OHG	OE	ON	Goth.	OHG	OE	ON
sg. N	<i>dags</i>	<i>tag</i>	<i>dæg</i>	<i>dagr</i>	<i>giba</i>	<i>geba</i>	<i>gifu</i>	<i>skor</i>
V	<i>dag</i>	<i>tag</i>	<i>dæg</i>	<i>dagr</i>				
A	<i>dag</i>	<i>tag</i>	<i>dæg</i>	<i>dag</i>	<i>giba</i>	<i>geba</i>	<i>gife</i>	<i>skor</i>
D	<i>daga</i>	<i>tage</i>	<i>dæge</i>	<i>dege</i>	<i>gibai</i>	<i>gebu</i>	<i>gife</i>	<i>skor</i>
G	<i>dagis</i>	<i>tages</i>	<i>dæges</i>	<i>dags</i>	<i>gibos</i>	<i>geba</i>	<i>gife</i>	<i>skarar</i>
I		<i>tagu</i>	<i>dæge</i>					
pl. NV	<i>dagos</i>	<i>taga</i>	<i>dagas</i>	<i>dagar</i>	<i>gibos</i>	<i>gebā</i>	<i>gifa</i>	<i>skarar</i>
A	<i>dagans</i>	<i>taga</i>	<i>dagas</i>	<i>daga</i>	<i>gibos</i>	<i>gebā</i>	<i>gifa</i>	<i>skarar</i>
D	<i>dagam</i>	<i>tagum</i>	<i>dagum</i>	<i>dōgom</i>	<i>gibom</i>	<i>gebōm</i>	<i>gifum</i>	<i>skōrom</i>
G	<i>dage</i>	<i>tago</i>	<i>daga</i>	<i>daga</i>	<i>gibo</i>	<i>gebōno</i>	<i>gifena</i>	<i>skara</i>

15.18. Of all the endings represented above, Modern English preserves only two of them – the *a*-stem genitive singular, continued today as the possessive ending *'s*, and the *a*-stem nominative plural, today our plural *-s*. The genitive, Common Germanic **-as(a)*, comes from PIE **-os(ǵ)o* (§6.48), while the nominative plural, Common Germanic **-ās*, comes from PIE **-ōs* (§6.52).

ADJECTIVES

15.19. **Strong and weak adjectives.** Germanic developed two different adjective declensions, called strong and weak. The so-called strong declension has endings modeled on those of the demonstrative pronouns: German *mit frisch-em Wasser* 'with fresh (dat.) water', with the same ending as the demonstrative pronoun *dem* 'the, this' (dat.). The weak declension has incorporated extra suffixal material in the oblique cases that comes from the PIE individualizing suffix **-on-*, which was added to adjectives to form words meaning 'the one that is X'. (It is also found in Latin names like *Catō* [stem *Catōn-*] 'the one who is *catus*, clever'.) Weak adjectives occur chiefly after a determiner or pronominal adjective, as in German *mit dem frisch-en Wasser* 'with the fresh water'. (Compare the similar phenomenon in Balto-Slavic of the "definite adjectives" discussed in §8.28 and §18.13.)

15.20. **Comparison of adjectives.** Germanic inherited the weak stem **-is-* of the IE comparative suffix **-jos-* (§6.78). Because the accent followed the suffix in PIE, the

suffix underwent Verner to become *-iz-, as in Goth. *manag-iza* 'more' (from *manags* 'many'). The Germanic superlative suffix was *-ista-, as in Goth. *manag-ista* 'most'. This goes back to PIE *-is- further suffixed with *-to-, just like -isto- in Greek (§6.81).

The normal English comparative and superlative suffixes -er and -est actually do not go back to the suffixes above, but to related forms that were innovated within Germanic, *-ōz- and *-ōsta-: Gothic *frod-ōza* 'wiser' and *arm-ōsta* 'poorest', and OE *liof-ora* 'dearer' and *liof-ost* 'dearest'. These suffixes became more common than *-iz-/ista- in North and West Germanic.

15.21. Pronouns. The first and second personal pronouns retain the dual in all the older Germanic languages, a category otherwise lost except in the Gothic verb: OE *ic* 'I', *wit* 'both of us', *wē* 'we'; þū 'thou', *git* 'both of you', *gē* 'you (pl.)'. The PIE demonstrative pronoun *so *seh₂ *tod 'this, the' is nicely preserved in Germanic, e.g. Goth. *sa so þata*, OE *sē sēo þæt*.

VERBS

15.22. The Germanic verbal system is much simpler than it was in PIE. Only the present and the perfect stems have remained, although some aorist inflectional endings were incorporated into the perfect. The perfect became a simple past tense called the preterite, but for the most part without reduplication. Also lost were the PIE imperfect, the subjunctive, and the future or desiderative formations. The optative survived, and became the Germanic subjunctive (the development and terminology are the same as in Italic, §13.19). The mediopassive survived as a living category to some extent in Gothic. Germanic has kept the present participle in *-nt- (§5.56, and see further below) and the verbal adjectives in *-tō- and *-onō- (a variant of *-nō-, §5.57).

15.23. Strong and weak verbs. Germanic verbs belong to two broad classes, termed strong and weak, according to how they make their past-tense forms (preterite and past participle). The distinction has its origin in PIE. Germanic (or pre-Germanic) verbs that could form perfects used the perfect as their past tense; these are the strong verbs. They show ablaut of the root in the formation of the tense-stems and form their past participles with the nasal suffix *-onō-. Strong verbs include most of the present-day "irregular" verbs like *sing sang sung* or German *singen sang gesungen*. By contrast, those verbs that did not form perfects (typically secondary verbs such as denominatives and causatives) developed an entirely new kind of past tense with an innovated suffix in *-d-. These are the weak verbs, and they typically do not ablaut. The great bulk of verbs in any Germanic language are weak (the "regular" verbs in today's grammar), such as Eng. *settle settled settled* or its German equivalent *siedeln siedelte gesiedelt*.

15.24. Preterite of strong verbs. While all Germanic verbs have a present, preterite, and past participial stem, strong verbs further distinguish between a preterite singular and plural stem. This faithfully reflects the distinction between the singular and plural stems of the PIE perfect (§5.47). Examples of Germanic singular/plural preterite pairs include OHG *bant* 'he bound' ~ *buntun* 'they bound', ON *greip* 'he grasped' ~ *gripo* 'they grasped', and Goth. *warþ* 'he became' ~ *waúrþun* 'they became' (where *au* represents etymologically the vowel *u*). The vowel *a* that is typical of the singular stems (*bant*, *warþ*; *greip* is from **graip*) comes from the **o* of the *o*-grade of the PIE perfect singular stem; and the zero-grade of the PIE perfect plural is reflected in *buntun*, *gripo*, and *waúrþun*.

15.25. In many preterites of this kind there is also a change in the stem-final consonant from the singular to the plural, as in OE sing. *wearþ* 'he became' but pl. *wurð-on* 'they became'. These consonantal changes are due to Verner's Law. Recall that the PIE perfect shifted the accent in the plural from the root to the endings (§5.47); in the plural, therefore, the accent came after the root-final consonant. If the root-final consonant was a voiceless stop, it underwent Verner's Law. Thus:

PIE perfect (without reduplication)	Grimm I	Verner	Later changes
3rd sing. <i>*uórt-e</i> 'he turned'	<i>*wórtþe</i>	—	<i>*warþ</i>
3rd pl. <i>*uŕt-ér</i> 'they turned'	<i>*wŕþér</i>	<i>*wŕðér</i>	<i>*wurð-un(t)</i>

(The original 3rd pl. ending **-ér* was replaced at some point by **-un(t)* from **-ŕt*, an aorist ending.) The Vernerized form of the root in the preterite plural is also found in the past participle, since it likewise goes back to a PIE form with accent on the suffix: **uŕt-on-ó-* '(having been) turned' > OE *wurden* 'having become'.

Forms like OE *wearþ* ~ *wurdon/wurden* are called Verner's variants, exhibiting what the Germans call *grammatischer Wechsel* or 'grammatical change'. They are abundantly preserved in Old English, Old High German, Old Saxon, and Old Norse. Interestingly, they are quite rare in Gothic, which leveled out the consonantal differences in almost all cases. The same development eventually overtook the other Germanic languages. Modern English is the only Germanic language that still preserves one pair of Verner's variants within a single tense-paradigm: the preterites *was* (sing.) and *were* (pl.), from OE *wæs wæron* (with *r* the English outcome of **z*, the Vernerized **s*; see §15.47).

15.26. Strong verb classes. Strong verbs are divided into seven classes, of which only the first three show the ablaut relationships inherited from PIE with perfect clarity. The modern English strong verbs have undergone so many changes that it is difficult to use them as examples of the original ablaut relationships, so the examples below are taken from Gothic. The forms given are the present infinitive, 1st sing. preterite, 1st pl. preterite, and the past participle of the verbs meaning 'bite', 'enjoy', 'bind', and 'throw'.

Class	PIE pattern	Germanic pattern	Gothic example			
I	<i>ei-oi-i-i</i>	<i>i-ai-i-i</i>	<i>beitan</i>	<i>bait</i>	<i>bitum</i>	<i>bitans</i>
II	<i>eu-on-u-u</i>	<i>eu-au-u-u</i>	<i>nīutan</i>	<i>naut</i>	<i>nutum</i>	<i>nutans</i>
III	<i>eRC-oRC-RC-RC</i>	<i>eRC-aRC-uRC-uRC</i>	<i>bindan</i>	<i>bamd</i>	<i>bundum</i>	<i>bundans</i>
			<i>waírpan</i>	<i>warþ</i>	<i>waúrpum</i>	<i>waúrpans</i>

As can be seen, from the PIE point of view these all behave identically, with the present stem continuing the full grade, the preterite singular continuing the o-grade, and the other forms continuing the zero-grade. In Class I, PIE **ei* became **i* in Germanic, preserved as such in Gothic but written *ei*. In Class III, Gothic *ai* and *au* represent short vowels from **e* and **u*.

The other classes do not reflect PIE ablaut as directly. Class IV, from IE roots ending with a single nasal or liquid, has replaced the zero-grade of the preterite plural with a long *ē*, e.g. Gothic *baíran bar berum baúrans* 'bear' (*e* writes long *ē* in Gothic). The same is true of Class V (from IE roots ending in a stop or *s*), which has additionally changed the past participle to full grade, e.g. Gothic *giban gab gebum*

gibans 'give'. Class VI is even more distant from the IE state of affairs, reflecting a Proto-Germanic pattern *a-ō-ō-a*: Gothic *faran* for *forum farans* 'go, fare' (o represents long *ō* in Gothic). This pattern may be related to that seen in Latin *scabō* 'I scrape', perfect *scabī* 'I scraped' (cf. §5.46). Finally, Class VII, substantially represented only in Gothic, contains a panoply of verbs whose preterites show reduplication, such as *haitan haihait haihaitum haitans* 'name' (where again *ai* represents a short vowel from **e* and *ai* a true diphthong). Only traces of reduplicated preterites are found outside Gothic, e.g. OE *hehton* 'they were named', ON *vera* '(he) rowed'.

15.27. Weak verbs. Weak verbs formed their past-tense forms (preterite and past participle) by the addition of a morpheme in *-*d-*, called the **dental preterite**, whose modern English descendant is the suffix spelled *-(e)d* or *-t*. The origin of the suffix is obscure, but it is usually assumed to be from the same root as the verb *do*. Weak verbs comprise mostly non-ablauting verbs, as discussed above. There were three classes of these verbs that can be established for Common Germanic. Class I contains denominatives and causatives formed with the suffix *-*ja-* (< PIE suffixes *-*ǵelo-* and *-*éǵelo-*). Examples include the Goth. denominative *andbahtjan* 'serve', pret. *andbahtida* (from the noun *andbahti* 'service, office'), and the causative *waljan* 'choose', pret. *walida*. Class II contains denominative verbs formed with a different suffix, *-*ō-* (from *-*āǵelo-*), e.g. Goth. *salbon* 'anoint', pret. *salboda*. Class III contains mostly stative and durative verbs like OHG *habēm* 'I have', Goth. pret. *habaida* 'had'. (A fourth class with a nasal suffix, represented e.g. by Goth. *fullnan* 'become full', are of uncertain status for Common Germanic.)

The past participle of weak verbs uses a suffix descended from the PIE verbal adjective suffix *-*tó-*. In general this was Vernerized to *-*da-*, as in Goth. *salbod-* 'anointed'; but the *-t-* remained in some consonant clusters, as still found e.g. in Eng. *brought* < Germanic **brāhta-*.

15.28. Other verbs. Four root athematic verbs survive, of which only one, 'be' (PIE **h₁es-*), is found in all of the family: Goth. singular paradigm *im is ist*, ON *em est es* (later *er*), and OE *eom eart is*. In West Germanic, 'do' is also athematic, as in OS *dōm dōs dōt*. Mention may also be made of the irregular verb *will*, which is historically an optative of a root present from PIE **uel-* 'wish' (cp. §15.10).

15.29. Preterite-presents. Several perfects/preterites have retained present-tense meaning (on this aspect of the perfect in PIE, see §5.49), and from them new preterites and participles were formed with the dental preterite suffix. In English, these primarily survive as the modal or "helping" verbs *can*, *shall*, *may*, and a few other forms. Since the present tense of these verbs is morphologically identical to a strong-verb preterite (note the lack of *-s* in the 3rd singular in Modern English), they are called preterite-presents or preterito-presents.

15.30. Subjunctive. As noted above, the Germanic subjunctive continues the PIE optative. Two tenses are distinguished, a present and a past. In Modern English, the subjunctive is formally distinguished from the indicative usually only in the third singular. Present subjunctives are found in such constructions as *Long live the Queen*, *Be that as it may*, and *(I asked) that he go*; past subjunctives (distinguished from past indicatives only in the verb *be*) are found for instance in contrafactual statements like *If he were there (I would go)*. Since the IE optative had secondary endings, the 3rd sing. ended in *-*t*, which disappeared in Germanic (§15.10 above); that is why 3rd singular subjunctives like *live* and *go* above seem to be "missing" the 3rd singular ending. (Analogously in German, the subjunctive *lebe* in *Es lebe die Königin* '[Long] live the queen' lacks the 3rd person *-t* of the indicative *leb-t* 'live-s'.)

15.31. Participles. The PIE present participle in **-nt-* survives as *-nd-*: Goth. *bairands* ‘bearing’, OE *berende*, OHG *beranti*, ON *berande*. Later in its history English replaced this with the unrelated suffix *-ing*, but a few traces of the old *nd-* participles can still be found in fossilized forms such as *friend* (originally a present participle **frijōnd-* ‘loving’) and *fiend* (Gmc. **fijand-* ‘hating’). The past participles, from PIE **-tō-* and **-onō-*, have already been discussed.

Syntax

15.32. The historical and comparative syntax of both the ancient and the modern Germanic languages has been the topic of much discussion by theoretical syntacticians. One of the issues of interest has been the phenomenon often called “V2,” or verb-second position. The phenomenon is most easily seen in German, which requires that all finite verbs in main clauses immediately follow the first syntactic constituent of the clause. This feature is found to some extent in Modern English, particularly with negatives: *Never did she look so gorgeous; Hardly had I entered the room when . . .* However, in the older stages of these languages, the word order tended to be freer.

It is also characteristic of several of the Germanic languages that in embedded (subordinate) clauses, the position of the finite verb is different from its position in main clauses. Thus in German one says *Ich habe ihm das Buch gegeben* ‘I have given him the book’, but (*Es ist wahr,*) *daß ich ihm das Buch gegeben habe* ‘(It is true) that I have given him the book’: in the second sentence the finite verb *habe* ‘have’ falls at the end of the embedded clause beginning with *daß* ‘that’. This was the preferred order in subordinate clauses in all the older Germanic languages.

The study of Germanic syntax is somewhat hampered by the scarcity of useful material from Gothic: Wulfila’s translation of the Bible (see §15.38 below) sticks very close to the word order of the Greek original. However, Gothic does preserve an archaic feature of PIE syntax not found elsewhere in Germanic, namely the ability to place clitics between preverbs and verbs (see §15.43).

Runic

Runes

15.33. The first writing system used by Germanic peoples to record their own languages is called the runic alphabet, after the *runes*, the name for the early Germanic letters. This word has been taken over into Modern English from OE or ON *rūn* ‘secret, mystery, written message’. From the available evidence it appears that runes were originally restricted to use in magic or religious rituals. The alphabet most likely stems from a similar alphabet that developed in northern Italy around the mid-first millennium BC and is an offshoot of a northern Etruscan alphabet. An early transitional version of the runic alphabet was used to write the Germanic name *harigasti* (also read *hariḡasti*; the name means ‘army-guest’) on a helmet unearthed in Negau, a Slovenian town near the Austrian border. The age of the inscription is disputed, but could date anywhere from the third century BC to the beginning of the first century AD; it is at any rate the earliest known inscriptional Germanic. All the other early Germanic runes are from much more northerly regions. The oldest so far discovered

are inscribed on a fibula (brooch) from the town of Meldorf in western Schleswig-Holstein, Germany, from the mid-first century AD, as well as some inscriptions from around AD 200 found in the drained valley of the river Illerup near Århus in eastern Denmark. These inscriptions are all quite short, but are not without linguistic interest. More extensive inscriptions begin to appear in the fourth century.

The Germanic runic alphabet is called the *futhark* (variously spelled), after its first six letters *f u þ a r k* (cp. the word *alphabet* after the first two letters *alpha* and *bēta* in Greek). The idiosyncratic angular shape of the runic letters, which eschew both curves and right angles, may be due to the fact that they were designed for inscribing on wood: all strokes had to run at an angle to the grain to keep the wood from splitting, straight lines are easier to carve in wood than curved ones, and strokes that cross the grain are easier to see.

15.34. Runes were used for all the Germanic languages, although by far the most numerous runic remains come from Scandinavia, particularly Denmark and Sweden. The language of the earliest runes, up till around AD 500, is rather uniform across a fairly wide geographical expanse, and has been named Runic or Runic Norse. Its position within Germanic is controversial. It is very similar to the putative ancestor of both North and West Germanic. However, the fact that it still preserves final *-z* (see next section) probably tips the balance in favor of specifically North Germanic filiation: North Germanic preserved *-z* but West Germanic lost it, and it is unlikely that West Germanic had not yet lost its **-z* by the Runic period. Nonetheless, we treat Runic separately here from the rest of North Germanic because of its age and controversial filiation.

15.35. Runic is important in preserving final *-z* before its later rhotacization to *-r* by the time it became Old Norse, as in *gastiz* 'guest' (ON *gestr*). The rune for this sound is frequently transcribed *ᚾ*, on the assumption that its phonetic value was between that of a *z* and an *r*; but this assumption is unnecessary. Runic also preserves short vowels before final *-s* or *-z*, as shown also by *gastiz*; these were later lost.

Most texts in Runic, before it differentiated into various local varieties that can be termed Old Swedish, Old Danish, etc., are rather short, telling little more than the name of the owner or manufacturer of the object on which the inscription appears.

Runic text sample

15.36. Inscription on one of the two large horns found in Gallehus in southern Jutland, Denmark, dating to c. AD 400. The first two words are written together. The inscription is poetic; like later Germanic poetic lines, it consists of two half-lines connected by alliteration (*ek hlewagastiz holtijaz || horna tawido*).

ekhlewagastiz . holtijaz . horna . tawido .

I, Hlewagastiz Holtijaz, made (this) horn.

15.36a. Notes. *ek*: 'I', Runic stressed 1st person pronoun, which also appears as suffixed (unstressed) *-ka*; elsewhere in Germanic the form is *ik*. *hlewagastiz*: personal name, 'famous guest'; the elements of this name are found in many other IE personal names (cp. §§2.48, 6.82). *Hlewa-* is from **kleyu-*, cp. Gk. *klé(w)os* 'fame'. *holtijaz*: 'of Holt(i)', meaning either 'son of Holt(i)' or 'from a place called Holt(i)'. *tawido*: 'made', 1st sing. weak preterite of a verb cognate with Goth. *tawjan*, preterite *tawida*.

East Germanic

15.37. East Germanic is the one extinct branch of Germanic and is represented principally by Gothic. According to the traditional history told by the sixth-century Gothic historian **Jordanes**, the **Goths'** ancient homeland was in Scandinavia; linguistically this may be supported by the name of the island of Gotland off eastern Sweden and other place names apparently containing *Got(h)* or a related element. (The present speech on Gotland, called Gutnish, is different from Swedish in certain ways and has lexical parallels with Gothic; for this reason some researchers have argued that Gutnish is closer to Gothic than it is to anything else, but this is not likely.) In or slightly before the second century AD, according to Jordanes, a group of Goths migrated from Scandinavia to the Baltic coast, defeating the Vandals and other Germanic tribes who lived there, and over several generations migrated farther south and east, arriving ultimately at the Black Sea. The Goths who settled in what is now Ukraine were the East Goths or **Ostrogoths**, while their relatives to the west,



Map 15.1 The Germanic peoples around AD 500

in what is now Romania between the lower Danube and Dniestr Rivers, are called the West Goths or Visigoths (*Visi-*, also written *Vesi-* in classical sources, actually does not mean 'west', but probably 'good, noble', from PIE **ues-*; cp. Ved. *vásu-* 'good').

The Visigoths were attacked by Huns in the year 376, who forced them across the Danube into Roman territory. For the next half-century they wandered through various parts of Europe in search of a place to live, along the way becoming converted to Arian Christianity, sacking Rome in 410, and finally settling in southern Gaul and Spain, where they founded a kingdom. After reaching its height in the late sixth and early seventh centuries, it was ultimately defeated and absorbed by the Muslim invaders of Spain in 711. As for the Ostrogoths, for a time they were also defeated by the Huns, but freed themselves in the fifth century and took over the rule of Italy until the mid-sixth century, which became the western part of a large kingdom stretching as far east as the Don River. We have a number of Ostrogothic names preserved in classical sources from this time.

One group that had settled in the Crimea retained their identity as Goths perhaps as late as the eighteenth century. Several dozen words and phrases in the language, termed Crimean Gothic, were collected in the early 1560s by a Flemish nobleman, Ogier Ghiselin von Busbecq, from a Greek in Constantinople. The transmission of these words seems to be inaccurate and corrupt in places, but is a precious window on a late type of Gothic different in certain ways from the dialect of the Gothic Bible (see below).

Also driven westward by the incursion of the Huns were the Vandals, whose language, Vandalic, was East Germanic as well. Among their various conquests, the Vandals overran part of Spain in the fourth century AD, whence the name of the historical region called *Andalusia*. Other East Germanic languages were Burgundian, Gepidic, and Rugian. All of these have essentially disappeared without a trace, and are known only from scattered personal and place-names.

Gothic

15.38. Our knowledge of Gothic stems almost entirely from the remains of a translation of the New Testament by a West Gothic bishop named Wulfila or Ulfilas ('Little Wolf', c. 311–382), which he undertook for the Visigoths living along the lower Danube. In addition, we have a few fragments of the *Skeireins* ('explanation, exegesis', a commentary on the Gospel of St. John and probably not written by Wulfila), and some isolated inscriptions and words preserved in other texts.

15.39. The spelling and alphabet were invented by Wulfila but are based on Greek. In transcriptions of Gothic words, *q* stands for the labiovelar *kʷ* and *hʷ* stands for the voiceless labiovelar glide *hʷ* (as in some pronunciations of English *what*). The letter *g* has the value [ŋ] before another *g*, *k*, or *q* (thus *siggwan* 'sing', *sigqan* 'sink'), but *gg* can also represent a double *g* (especially when it arose by *Verschärfung*, see §15.41 below). The vowels are slightly complicated. Wulfila used *ai* and *au* each in two different values, as the diphthongs *ai* and *au* and as the short close vowels [ɛ] (like in Eng. *bet*) and [ɔ] (like in *bought*); to distinguish these, the transcription *ai* and *au* is frequently used for the short-vowel values (we shall do this here). The vowels *e* and *o* stand for long open *ē* and *ō* ([e:] and [o:]), as in Eng. *bare* and *bore*). The letter *u* stands for both short and long *u*, and *ei* stands for long [i:] (like Eng. *beat*).

Phonology

15.40. Although Gothic is the oldest Germanic language preserved in any abundance, it is not in every respect the most conservative or archaic; as with any language, it contains its own unique mixture of old and new. Most noteworthy in this regard is the fact that it had gotten rid of nearly all the Verner's variants (§15.25) by the time it is attested, by generalizing one or the other stem-final consonantal variant throughout any given paradigm. For example, the verb *wairþan* 'become' has singular preterite *warþ* as expected, but plural *wairþum* rather than expected **wairdum* (contrast OE *wearþ*, *wurdon*). On the other hand, Gothic preserves the Proto-Germanic vowel inventory better than any other Germanic language. For one thing, it was not affected by any of the various umlauts that spread through North and West Germanic territory some centuries later (§§15.48 and 15.89–91).

15.41. Another phonological feature of Gothic, this time shared with North Germanic, is the *Verschärfung* (or "hardening") of the Proto-Germanic geminate glides **jj* and **ww* into geminate stops plus glide. The geminate glides arose from diverse sources that need not be entered into here. In Old Norse, **ww* and **jj* became *ggv* and *ggj*, while in Gothic the outcomes were *ggw* and *ddj*. For example, from Gmc. **treww(i)a-* 'trustworthy, true' we have Goth. *triggws* and ON *tryggva*, and from **twajjē(n)* or **twajjō(n)* 'of the two' (genit. pl.) we have Goth. *twaddje* and ON *tvægja* (cp. OHG *zweio*). Crimean Gothic has simple *d* here, as in *ada* 'egg' from Gmc. **ajja-*; compare ON *egg* (from which Eng. *egg* was borrowed).

Morphology

15.42. Gothic preserves some important morphological archaisms not found in the other Germanic languages. Verbs have separate forms for the first and second persons dual, and a passive conjugation in the present tense. (The passive in Germanic is otherwise found only in the 1st singular of the verb 'to be called, named': Runic *haite* 'I am called', ON *heiti*, OE *hätte*, Mod. German *ich heiße*.) A vocative different from the nominative is preserved in some declensions (e.g. nomin. *dags* 'day', voc. *dag*). On the other hand, Gothic has lost the instrumental, which does survive in some of the other languages.

Syntax

15.43. Gothic preserves several enclitic particles that occur in the second position of a sentence by Wackernagel's Law (§§8.22ff.). These occur in chains, just as in other older IE languages (cp. Anatolian, §9.13), and preverbs count for the purposes of determining their placement – that is, second-position clitics will intervene between a preverb and its associated verb if the preverb and verb stand at the beginning of their clause. For example, the interrogative particle *u* appears between the preverb and verb in *ga-u-laubjats* 'Do you (both) believe . . . ?' (Matt. 9:28); the verb is *ga-laubjan* 'to believe', cognate with OHG *gilouban* (Mod. German *glauben*) and (with different prefix) Eng. *be-lieve*. An example of a chain consisting of the same particle plus the indefinite object pronoun *hwa* 'something' is seen in the sequence *ga-u-hwa-sehvi* 'whether he saw anything' (Mark 8:23). All of this is quite foreign to North and West Germanic.

15.44. Not unexpectedly in light of these facts, Gothic is the only Germanic language to preserve a living descendant of the PIE enclitic conjunction **k^we* 'and' (§7.26), which shows up as *-h* (e.g. *ga-h-melida* 'and he wrote') or *-uh* (IE **u-k^we*, e.g. *urreis nim-uh* 'arise and take!').

Gothic text sample

15.45. Mark 8:14–18. For the orthography, see §15.39 above.

14 jah ufarmunnodedun niman hlaibans jah niba ainana hlaif ni habaidedun miþ sis in skipa. 15 jah anabaup im qipands: saifviþ ei atsaiþiþ izwis þis beistis Fareisaie jah beistis Herodis. 16 jah þāhtedun miþ sis misso qipandans: unte hlaibans ni habam. 17 jah fraþjands Iesus qap du im: hva þaggkeiþ unte hlaibans ni habaiþ? ni naúh fraþiþ nih wituþ, unte daubata habaiþ hairto izwar. 18 augona habandans ni gasaiþiþ, jah ausona habandans ni gahauseiþ jah ni gamunup.

14 Now they had forgotten to bring bread; and they had only one loaf with them in the boat. 15 And he cautioned them, saying, "Take heed, beware of the leaven of the Pharisees and the leaven of Herod." 16 And they discussed it with one another, saying, "We have no bread." 17 And being aware of it, Jesus said to them, "Why do you discuss the fact that you have no bread? Do you not yet perceive or understand? Are your hearts hardened? 18 Having eyes do you not see, and having ears do you not hear? And do you not remember?"

15.45a. Notes. 14. jah: 'and'; the *-h* is from PIE **k^we* 'and'. ufarmunnodedun: 'they forgot'; -dedun is the 3rd pl. weak preterite ending, and ufar- is a preverb, lit. 'over' < **super* (Gk. *hupér*, Lat. *s-uper*). niman: 'to take', German *nehmen*, OE *niman*; PIE **nem-*. The past participle of the OE verb lives on today as *numb* (§15.13). hlaibans: 'loaves of bread', cognate with Eng. *loaf* (OE *hlāf*); -ans is accus. pl., preserving the nasal (§6.55). niba: 'not even', combined with the following *ni* in a double negative construction. habaidedun: *hab-* is the root, like German *haben* 'to have'; the rest is stem plus weak dental preterite. miþ: 'with', cognate with German *mit* and the mid- of Eng. *mid-wife* (literally '[one who is] with the woman [giving birth]'). skipa: 'boat', cognate with both *ship* (the native English word) and *skiff* (ultimately from a Germanic language that preserved *sk-* like Gothic).

15–16. im: 'them', dat. pl. qipands: 'saying'; -ands is the nominative of the present participle, and qipan is cognate with archaic English *quoth* (past tense, equivalent to Goth. *qap* as in verse 17 below). saifviþ ei: 'see (2nd pl.) (to it) that . . .'. Sathuan, cognate with Eng. *see*, German *sehen*, is from **sek^w-*; the 2nd pl. ending -iþ is the same as ON -ið, German -(e)t (imperative *seht!* 'see ye!'), from PIE **-ete*. izwis: 'you (pl.)', dat. and accus.; the genit. is *izwar*, seen below, verse 17. þis beistis: 'the leaven', genit. sing. object of the verb *atsaiþiþ* 'you look out for'. þāhtedun: 'thought', Gmc. **þāhtō* from earlier **þanbtō*, with loss of nasal as per §15.11; the nasal is still in the present stem (*þagkjan*, Eng. *think*). misso: '(each) other'. unte: cognate with German *und*, Eng. *and*, but here meaning 'that' after the verb of saying.

17. fraþjands: 'understanding, being aware'. du: 'to', with unexplained *d-*. hva: lit. 'what', here meaning 'why'; from PIE **k^wod*, with regular loss of the final *-d* (preserved in the demonstrative *pata* 'that (thing)' by the addition of some enclitic particle). naúh: 'yet', = German *noch* 'yet, still', IE **nu-k^we*. nih: 'nor, and not', IE **ne-k^we* (> Lat. *neque* 'nor'). wituþ: 'you (pl.) know'; the singular stem is *wait-*, from the old PIE perfect **uoid-*, **uid-*, see §5.48. daubata: 'deaf, hardened', neut. accus. of the adj. *daufs*; -ata is the 'strong' adjectival ending, taken over from pronominal forms like neut. accus. *pata* 'that'. hairto: 'heart', a neuter *n*-stem, hence its inflection in German as *Herz* (nomin.-accus.) but dat. *Herz-en*.

18. *augona*: 'eyes', neut. nomin.-accus. pl., also a neuter *n*-stem (German pl. *Auge-n*), from Gmc. **augōn*, ultimately from PIE **h₂ek^w-* (Lat. *oc-ulus*, OCS *oko*, etc.) but influenced by Gmc. **auzōn* 'ear' (PIE **h₂eus-*; the Gothic descendant, in the neut. pl., is *ausona*, a few words later in this verse). *gasaiſviþ*: 'see', 2nd pl., with the addition of the prefix *ga-* to give it a punctual or perfective sense. The same prefix is found in *gahauseiþ* 'you hear' and *gamunniþ* 'you understand'; the Greek original has ordinary presents for all of these, so this is clearly a Gothic feature and not a literal rendition of the Greek. *gahauseiþ*: 'you hear', from the verb *hausjan* 'to hear'. In the rest of Germanic it shows the effects of Verner's Law (Eng. *hear*, ON *heyra*, all from **hauzja-*). The PIE root for 'hear' is **h₂kous-*, as in Gk. *akoúō* 'I hear'.

West Germanic

15.46. Three ancient West Germanic languages are attested in meaningful quantities from before the year 1000: Old English, Old Saxon, and Old High German. Old High German is more divergent from the other two than either of them are from each other (Old Saxon and Old English are quite similar and must have been mutually intelligible). For this reason, a division is often made between Old High German and the rest of West Germanic.

Consonants

15.47. In Common West Germanic, Germanic final **-z* was lost, as in OE *dæg* and German *Tag* 'day' from **dagaz* (cp. Goth. *dags*, ON *dagr*). Elsewhere, **z* was rhotacized to *r*, as in **luzana-* 'lost' > Eng. *lorn* (as in *love-lorn*, *for-lorn*). A consonant plus the glide **j* became a geminated consonant following a short vowel, as in many verbs with the suffix **-ja-* (see §15.27), e.g. **sitjan* 'to sit' > OE *sittan*, **bidjan* 'to ask' > OE *biddan* (cp. Goth. *bidjan*). Additionally, *h* was lost before consonants word-initially; this happened independently in all the West Germanic languages in the Middle Ages. Thus contrast Eng. *loud* and German *laut* with OE *hlūd* and OHG *hlūt*, *hlūd*. West Germanic loanwords beginning with *hl-* borrowed into early French, which had no *h*, were borrowed as *fl-* or *cl(h)l-*, as in the name of the Merovingian king *Clovis* from Frankish **Hladowig* 'loud-battle'; a later form of the same name was borrowed as *Louis*.

Vowels

15.48. The Germanic vowels underwent more substantial changes in West Germanic than the consonants. Principal among these were the various *umlauts*, vowel shifts caused by the presence of a particular vowel or glide in a following syllable. By far the most significant, and most familiar, was *i*-umlaut. This sound change affected the back vowels *a*, *ā*, *ō*, *u*, and *ū* when they preceded a syllable containing *i* or *j* (the glide *j*). The back vowels then became fronted to vowels essentially equivalent to Modern German *ä*, *ö*, and *ü*. A tabular overview, with further remarks on other umlaut processes, is given below in the section on Old English, §15.57.

While *i*-umlaut ultimately affected most of the West Germanic-speaking area (as well as Old Norse, see §15.91 below), it did not affect all of it equally or simultaneously. It appears to have started in the northwest, as its effects are complete in

Old English from the earliest documented times. It also spread to Old Norse fairly early. Farther south, however, matters are less clear: umlaut is not indicated consistently in Old High German texts, and modern southern German dialects show it less consistently too.

Old English

15.49. Around the late fourth and continuing into the fifth and sixth centuries AD, three Germanic tribes inhabiting the coastal areas of Denmark and northern Germany began migrating to England: the Angles, Saxons, and Jutes. The traditional date for their arrival is 449, but there is good evidence that their migrations had already begun by 400. The Angles, inhabiting northern Schleswig-Holstein and already mentioned by the Roman historian Tacitus in the first century AD, settled large parts of Britain in the fifth and sixth centuries, including what would become the kingdoms of Northumbria, Mercia, and East and West Anglia. Around the early fifth century, the Saxons, who had aggressively expanded into northern Germany from their homeland in eastern Schleswig-Holstein, launched pirate raids against Britain and began to colonize it, settling in areas that are now called Essex, Wessex, and Sussex (literally 'East Saxons', 'West Saxons', 'South Saxons'). The Jutes came probably from Jutland in Denmark and settled in Kent in southeast England and the Isle of Wight. The original distinctions among these closely related peoples soon disappeared; medieval writers referred to the whole population simply as *Angli* or Angles, whence the OE name *Englaland*, 'land of the Angles, England'. In modern scholarship, they are collectively called the **Anglo-Saxons**.

15.50. The language of this new population in Britain is known as **Old English** or Anglo-Saxon. It is first known from fragmentary runic inscriptions of little linguistic value from the late fourth or early fifth century; the literary period did not start until several centuries later. The surviving corpus of Old English literature is substantial and consists of both poetry and prose. Most Old English poetry is anonymous and cannot be dated with certainty, but was probably written down between the eighth and eleventh centuries. The oldest datable poem is Caedmon's hymn (§15.66), composed around 650. The longest and most famous Old English poem is *Beowulf*, a heroic epic containing a mixture of historical and legendary material, some of it quite ancient.

The earliest Old English prose is older than the poetry. Its first known representative is the law code of King *Aethelberht* of Kent, written c. 600. The composition of literary prose prospered during the reign of King Alfred the Great (849–99) of Wessex, who instigated a cultural renaissance in the late ninth century and championed the use of Old English as a literary vehicle equal to Latin. Under Alfred's influence, the Wessex variety of the West Saxon dialect (spoken in southwestern England) became the literary standard, ushering in the period of "Classical" Old English. Literary activity in England was vigorous until the Norman Conquest (1066); many of the prose works produced during this time are in an excellent style on a par with that of any other contemporaneous literature in Europe.

15.51. Aside from West Saxon, Old English had three other major dialects: Northumbrian, spoken north of the Humber river, preserved in short inscriptions, poems, and glosses; Mercian, spoken between the Humber and the Thames, preserved in charters and glosses and the source of most Modern English native forms;

and Kentish, spoken in a small slice of southeast England, preserved also in charters and in some literature and glosses.

15.52. Old English shares several features with Frisian, leading many scholars to posit an Anglo-Frisian subgroup of West Germanic. The palatalization of *k* before front vowels (see the next section) and the “brightening” of *a* to *æ* (§15.57) are among the characteristics of Anglo-Frisian.

Developments of Old English

CONSONANTS

15.53. Palatalization. Most of the changes that Old English made to the West Germanic consonant system had to do with the velars. The voiceless velar stop **k* when neighboring front vowels (*e* and *i*) or preceding the palatal glide *j* was palatalized to the affricate [č] (written *c*, later *ch* starting in Middle English). This is the source of the *ch* in such Modern English words as *chew* (OE *cēowan*), *church* (OE *cirice*), *drench* (pre-OE **drankjan*), and *ditch* (OE *dīc*; contrast German *kauen*, *Kirche*, *tränken* ‘to water’, and Old Norse *díki* ‘ditch’, all without palatalization). The voiced velar stop **g* was weakened to the voiced velar fricative [ɣ] or the glide [j] (but still spelled *g* in Old English) in many positions, especially word-finally, between vowels, and before most front vowels. This is the source of the initial *y-* in such words as *yellow*, *yield*, and *yarn* (OE *geolu*, *geldan* ‘to pay’, *gearn*; compare German *gelb*, *gelten*, *Garn* with preserved *g-*); note also *day*, *many*, *saw* (‘saying’) from *dæg*, *manig*, and *sagu*. A hard *g* in these positions in a Modern English word typically marks it as a borrowing, as is the case with *get*, *guest*, *gift*, *egg*, *dreg*, and *nag*, which are all from Old Norse.

15.54. Other changes. A feature that characterizes not only Old English but all the West Germanic languages except High German is the prehistoric loss of nasals before fricatives (*f* *þ* *s*), with compensatory lengthening of the preceding vowel. Thus where OHG has *fimf* ‘five’, OE and Old Saxon have *fif*. Similarly, the word for ‘other’ in Gothic is *anþar* and in OHG *ander*, while OE has *ōþer*; and contrast also OHG *uns* ‘us’ with OE *ūs*, and German *Gans* with OE *gōs* ‘goose’. Additionally, the Germanic voiced fricative **ð* remained a fricative and turned into voiceless *f*, as in *hlāf* ‘loaf’ < **hlaibaz* (contrast German *Laib*).

15.55. Between vowels, the fricatives *f* and *s* were voiced to *ν* and *z* in Old English, though spelled as *f* and *s*: *frēosan* ‘freeze’, *ofer* ‘over’.

15.56. Finally, the consonant cluster **sk* became the sibilant [ʃ] in Old English (i.e. *sh*) and was spelled *sc*: *fisc* ‘fish’, *scip* ‘ship’, *wæscan* ‘wash’. Words in English with the sound *sk* are borrowings, often from other Germanic languages. Thus *shirt* is native English (OE *scyrte*) but *skirt* is a borrowing of its Old Norse cognate *skyrtu*. Similarly, the word *blatherskite* turns out to be a bit more colorful than its modern meaning (‘foolish person’) might suggest: its second element is a borrowing of the Old Norse cognate of English *shit*.

VOWELS AND DIPHTHONGS

15.57. The Germanic vowel system underwent more significant change than the consonants in OE. The back vowels **a* and **ā* became fronted (or “brightened”) in most positions to the front vowels *æ* and *ē* except before nasals: thus *ræt* ‘rat’

(vs. German *Ratte*, which preserved the old *a*) and *hæfde* 'had' (vs. German *hatte*). (One can see with this change part of the characteristic "sound" of English shining through at an early date.) The most far-reaching change was *i*-umlaut, causing the diphthongs and most vowels to be fronted before a syllable containing **i* or **j* in pre-Old English. The following table exhibits most of the changes; the first example for each vowel or diphthong is not umlauted, while the second example is umlauted and is boldfaced:

West Germanic		Old English	Modern English
* <i>a</i>	* <i>drank</i>	<i>dranc</i>	<i>drank</i>
	* <i>drankjan</i>	<i>drenčan</i>	<i>drench</i>
* <i>ō</i>	* <i>fōt</i>	<i>fōt</i>	<i>foot</i>
	* <i>fōti</i>	<i>fēt</i>	<i>feet</i>
* <i>u</i>	* <i>guda</i>	<i>god</i>	<i>god</i>
	* <i>gudiga-</i>	<i>gydig</i>	<i>giddy</i> (orig. 'possessed by god')
* <i>ū</i>	* <i>mūs</i>	<i>mūs</i>	<i>mouse</i>
	* <i>mūsi</i>	<i>mȳs</i>	<i>mice</i>
* <i>ai</i>	* <i>gaista</i>	<i>gāst</i>	<i>ghost</i>
	* <i>gaist(a)lik-</i>	<i>gæstlic</i>	<i>ghastly</i>
* <i>au</i>	* <i>staupa-</i>	<i>stēap</i>	<i>steep</i>
	* <i>staupila</i>	<i>stīepel</i>	<i>steeple</i>
* <i>eu</i>	* <i>þeuba</i>	<i>þēof</i>	<i>thief</i>
	* <i>þeubipō</i>	<i>þīeff</i>	<i>theft</i>

15.58. The new *æ* and *ǣ* from **a* and **ā*, together with the other front vowels *e*, *i*, and *ī*, underwent a diphthongization in various phonetic contexts known as *breaking*. Examples include the following:

Germanic	Old English	Modern English
* <i>liznōn</i>	<i>leornian</i>	<i>learn</i>
* <i>ahtau</i>	<i>eahta</i>	<i>eight</i>
* <i>erþō</i>	<i>eorþ</i>	<i>earth</i>

These diphthongs reverted to single vowels later in the history of English, but spellings like *learn* and *earth* are orthographic remnants of the earlier diphthongal pronunciation.

There were also diphthongs inherited from Germanic, and these underwent change as well. First, **ai* was monophthongized to *ā*, which eventually became long *ō* in most dialects, including the one ancestral to Modern English. This is still a diphthong in German, where it is usually written *ei*; hence English-German pairs like *bone/Bein*, *stone/Stein*, *home/Heim*, and *whole/heil*. The other two Germanic diphthongs, **au* and **eu*, became *ēa* and *ēo*: compare OE *cēap* 'bargain' (> modern *cheap*) with German *Kauf* 'sale', and OE *flēogan* 'to fly' with OHG *fliogan* (< Germanic **fleugan*).

15.59. Recognizing ablaut and umlaut in **Modern English**. English contains many words or word-pairs having vowel alternations; some are due to ablaut, some to umlaut, and telling them apart usually requires some knowledge of their history. Vowel changes in the principal parts of strong or "irregular" verbs are normally

vestiges of PIE ablaut; the same is true of many nouns derived from or related to them, such as *song* alongside *sing*, and *seat* alongside *sit*. Umlaut is the source of, among other things, the vowel alternations seen in irregular plurals (*mouse/mice*, *foot/feet*, *man/men*; these all go back typically to Germanic plurals with the suffix **-iz*), in abstract nouns ending in *-th* that are formed from adjectives (*long/length*, *foul/filth*, *whole/health*; the suffix was **-iþō-*, with the *-i-* causing umlaut), and in the occasional irregular comparative or superlative like *elder*, *eldest* alongside *old* (Germanic **aldiza-* **aldista-*; recall §15.20). In some cases, a vowel alternation in Modern English is due to a combination of both ablaut and umlaut, as with *stench* alongside *stink* (Germanic **stankja-* alongside **stinkan*) and with the causatives *drench* and (to) *fell* alongside *drink* and *fall*.

Morphology

15.60. The morphology of Old English was little changed from that of Common Germanic, and resembles that of Modern German. There were four cases in the noun and pronoun (nominative, genitive, dative, accusative), plus traces of an instrumental case; verbs were conjugated in a present and preterite indicative and in a present and past subjunctive, with different endings for most of the persons in singular and plural.

15.61. Many Old English nominal and verbal paradigms show the effects of umlaut; virtually all of these effects were leveled out in the later development of the language, with the occasional "irregular" plurals like *man/men* being essentially the only remnants. Examples of such paradigms follow: a root noun (*fōt* 'foot'), a consonant-stem (*fēond* 'enemy'), and an *s*-stem (*lomb* 'lamb'). Also, two strong verb paradigms are given in the present and preterite, an example of Class III (*helpan* 'to help') and of Class V (*cēosan* 'to choose', this one also showing Verner's variants in the preterite):

sg.				present			
N	<i>fōt</i>	<i>fēond</i>	<i>lomb</i>	sg. 1	<i>helpe</i>	<i>cēose</i>	
A	<i>fōt</i>	<i>fēond</i>	<i>lomb</i>	2	<i>hilpst</i>	<i>cīest</i>	
D	<i>fēt</i>	<i>fiend</i>	<i>lombe</i>	3	<i>hilpp</i>	<i>cīest</i>	
G	<i>fōtes</i>	<i>fēondes</i>	<i>lombes</i>	pl.	<i>helppaþ</i>	<i>cēosaþ</i>	
pl.				preterite			
N	<i>fēt</i>	<i>fiend</i>	<i>lombru</i>	sg. 1	<i>healp</i>	<i>cēas</i>	
A	<i>fēt</i>	<i>fiend</i>	<i>lombru</i>	2	<i>bulpe</i>	<i>cure</i>	
D	<i>fōtum</i>	<i>fēondum</i>	<i>lombrum</i>	3	<i>healp</i>	<i>cēas</i>	
G	<i>fōta</i>	<i>fēonda</i>	<i>lombra</i>	pl.	<i>hulpon</i>	<i>curon</i>	

It is instructive to note that Old English nouns had well over a half-dozen plural formations, of which the ancestor of the productive modern *s*-plural was merely one: *stān* 'stone' pl. *stānas*; *hēafod* 'head' pl. *hēafodu*; *word* 'word' pl. *word*; *giefu* 'gift' pl. *giefra*; *lār* 'teaching, lore' pl. *lāra*; *sāwol* 'soul' pl. *sāwla*; *eoh* 'horse' pl. *eos*; *mæd* 'mead, meadow' pl. *mædwa*; *oxa* 'ox' pl. *oxan*; and *here* 'army' pl. *herigeas*. Almost all of these have disappeared without a trace; *oxen* is a remnant of the old *n*-stem declension (note how what we think of as a plural ending *-en* was originally the stem, from the IE point of view), while *men* and *feet* go back to umlauting paradigms. The *-r-* in *children* (dialectally also *childer*) is the only remnant of the

lombru-type plural in the paradigm above, where the *-r*- ultimately continues the *-s*- of the IE *s*-stems. The developments by which *-s* became the default plural are complex and not fully documented.

In the verb paradigms, the old 3rd plural endings have been generalized throughout the plural (*-ap* < **-anþ* < primary **-onti*; *-on* < Germanic **-un* < **-nt*); Old English shares this characteristic with all of West Germanic outside of High German. The 2nd singular of the preterite, with its short-vowel ending and the same ablaut grade as the plural, is characteristic of West Germanic and unusual from the PIE point of view. The ending appears to go back to Gmc. **-īz*, the 2nd sing. optative ending; what we are probably dealing with is an old perfect optative that was used for politeness. (The normal 2nd sing. perfect ending **-thē* is preserved as *-t* in the preterite presents, as in early Modern Eng. (*thou*) *shalt*.)

Middle and Modern English

15.62. After the Old English period, almost all the inflections seen in the previous section were lost. This loss has been frequently but erroneously attributed to the influence of Norman French (on which more shortly). In fact, changes that predated the Norman period had already set the scene for it, in particular the loss of distinctions in the vowels of final unstressed syllables. In the Middle English period, these syllables were lost, and with them most inflectional endings. The different classes of Old English nouns were collapsed into essentially one class which added *-es* to form the possessive singular and the plural. (This had not fully affected feminine nouns or kinship terms in *-er*, as seen by phrases like *his lady grace* and *thi brother wife*.) Grammatical gender was also lost except in a few dialects, as well as the dual inflection in pronouns. To express grammatical roles of nouns, prepositional phrases and distinctions in word order became increasingly important.

15.63. Middle English lasted from the time of the Norman Conquest (1066) to about 1500. The first century and a half of this stage is poorly documented. The Normans not only ousted the Anglo-Saxon royalty and nobility, but also replaced the native clergy and scribes with speakers of Norman French. The forced break in the scribal tradition resulted in a precipitous drop in the writing of English, and by the thirteenth century, English had absorbed hundreds of French loanwords. But after the loss of Normandy in 1204, interest in French declined and the prestige of English rose again. London was the administrative capital and attracted a steady stream of settlers who brought regional dialect features with them. London English adopted some of these features, increasing its comprehensibility far from the mouth of the Thames; gradually it assumed the role of a standard language. By the end of the fourteenth century, English was the language of instruction in schools; it had also triumphed in the originally Norse-speaking Danelaw, the area of northern England settled by the Norsemen. (But Norse had left an indelible stamp on English vocabulary; see §15.85 below.)

15.64. The bulk of surviving Middle English literature consists of anonymous religious and didactic works in verse, including romances like the brilliant late fourteenth-century poem *Sir Gawain and the Green Knight*. A linguistically valuable piece of didactic religious verse, albeit devoid of literary interest, is the *Ormulum* by a canon named Orm (fl. 1200). This poem was written in a spelling system

devised by Orm himself to reflect the phonetic details of the language faithfully. Needing little introduction is the finest poet in English before Shakespeare, Geoffrey Chaucer (c. 1342–1400), who became a master of narrative poetry, especially with his unfinished *Canterbury Tales*.

15.65. The Modern English period includes English since 1500. Its chief linguistic development was the Great English Vowel Shift, a series of changes that had run their course by the eighteenth century (at least in standard English; some dialects were never affected by it, or affected only partially or in different ways). All the long low and mid vowels were raised, and the long high vowels became diphthongs consisting of [a] plus a glide:

Middle English		Modern English	
[i:]	<i>bīten</i>	[aj]	<i>bite</i>
[e:]	<i>sēen</i>	[ij]	<i>see</i>
[a:]	<i>gāte</i>	[ej]	<i>gate</i>
[o:]	<i>bōt</i>	[uw]	<i>boot</i>
[u:]	<i>lūd</i>	[aw]	<i>loud</i>

As can be seen from the Modern English examples, the older spelling was frequently maintained. Modern English spelling in fact typically reflects late Middle English pronunciation.

Old English text sample

15.66. Caedmon's Hymn, as preserved in the earliest Anglian (Northumbrian) version, AD 737; probably composed around 650. A *u* before a vowel represents *w*. The version on the right is in West Saxon.

Northumbrian

Nū scylun hergan heƿenricas Uard,

Metudæs maecti end his mōdgidanc,
uerc Uuldurfadur, suē hē uundra
gihuaes,

ēci Dryctin, ōr āstelidæ.

Hē ārist scōp aelda barnum
heben til hrōfe, hāleg Scepen;
thā middungeard moncynnæs Uard,
ēci Dryctin, æfter tīdæ,
firum foldu, Frēa allmectig.

West Saxon

Nū wē sculon herigean heofonrices
Weard,

Meotodes meakte ond his mōdgeþanc,
weorc Wuldorfæder, swā hē wundra
gehwæs

ēce Drihten, ōr onstealde.

5 Hē ārest sceōp eorðan bearnum
heofon tō hrōfe, hālig Scyppend;
þā middangeard monncynnes Weard,
ēce Drihten, æfter tēode,
firum foldan, Frēa ælmihtig.

Now we must praise the guardian of the heaven-kingdom,
the might of the Lord and his wisdom,
the work of the Glorious Father, how he,
eternal Lord, established the beginning of each of the miracles.

- 5 He first created for the children of men [var.: of the earth]
the heaven as a roof, the holy Creator;
then mankind's Guardian,
eternal Lord, made middle earth after that,
land for men, almighty Lord.

15.66a. Notes (selective). 1–3. *hefencīcaes*: 'of the heaven-kingdom', genit. sing. Northumbrian is valuable in distinguishing *æ* (written *ae* < **a*; see §15.57) from *e* in unaccented final syllables. The noun *ric* 'kingdom', Gmc. **rikja-*, is a loanword from Celtic (ultimately < PIE **h₂reǵ-* 'king', cp. Lat. *rēx* 'king', Ved. *rāj-an-* 'king'). *Metudæs*: 'Lord, God', genit. sing.; one of the native Germanic words adapted for use to refer to the Christian God. It originally meant 'fate' and comes from the verb *metian* 'to measure' (**med-*, cp. OIr. *midiur* 'I measure, estimate', Lat. *modus* 'measurement'). *mæcti*: 'might', feminine *ti*-abstract noun (§6.42) of the verb *magan* 'be able, be powerful' (> Mod. Eng. *may*). *mōdgīdanc*: 'wisdom', lit. 'mind-thought'; *mōd* (Mod. Eng. *mood*) is cognate with German *Mut* 'courage' and Goth. *moþs* 'anger', but of unclear etymology; *gidanc* is an example of a very common type of abstract noun in Germanic, made by adding the suffix **ga-* to a verbal root to form a collective. (The type survives in the noun *hand-iwork* < OE *hand-geweorc*.) uerc: 'work', PIE **uerǵ-* (Gk. (*w*)*érgon* and Arm. *gorc* 'work'). *uundra*: 'wonders', genit. pl.; the genit. pl. ending in OE was *-a*, from PIE **-dm*.

4–5. *ēci*: 'eternal', cognate with the first part of Goth. *ajuk-dūps* 'eternity' and ultimately derived from **h₂oju-* 'life(time)' via Gmc. **aiw-* (which became *ā(w)* in OE; cp. German *ewig* 'forever, eternal'). *Dryctin*: 'Lord', another Germanic term appropriated to refer to God; originally just 'lord (of battle)', cp. Goth. *ga-draht-s* 'soldier'. *ōr*: 'beginning', cognate with ON *óss* 'mouth (of a river)' and Lat. *ōs* 'mouth'. *āstelidæ*: 'established'; the preverb *ā-* is continued in Mod. Eng. *acknowledge* (from OE *ā-cnāwian* 'recognize, acknowledge'), and *stelidæ* is from *stellan* 'place, set up', cp. German *stellen*, ultimately from PIE **steh₂-* 'stand' with additional derivational material. *ārist*: 'first', derived from the superlative of Gmc. **air-* 'early, early in the day', from PIE **ajer-* 'morning, day' (Av. *aiara* 'day', Gk. *ār-iston* 'breakfast, first meal of the day'); the Germanic comparative **airiz* underlies Eng. *ear-ly*. *scōpi*: 'created', preterite of *sciepan* (cp. Mod. Eng. *shape* < OE *gesceap* 'a creation', and German *schaffen* 'create, accomplish'). *aelda*: 'of ages', i.e. 'of generations, of men', genit. pl.; a *ti*-abstract from the root **h₂el-* 'nourish' (also seen in OE *alan* 'nourish, raise', Lat. *al-umnus* 'one nourished', and Eng. *old*, lit. 'nourished, grown up' from the old *to*-participle). The West Saxon version has a different word here, *eorðan* 'earth'. *barum*: 'children', dat. pl.; the word survives in Scots *bairn*, and literally means 'one carried (in the womb)', from **bher-* 'carry'.

6–9. *hāleg*: 'holy', a derivative of *hāl* 'whole', from Gmc. **haila-* < PIE **kat-lo-* 'whole, complete' (cp. OCS *celi* 'whole', and the Old Prussian toast *Kails! Pats kails! Hail! Hail yourself!*, useful at parties). *middungeard*: 'middle earth', the Germanic term for the earth and humanity, and the inspiration behind J. R. R. Tolkien's fictional Middle Earth. *Middun* is ultimately from PIE **medhjo-* 'middle' (Lat. *medius* 'middle', Gk. *mésos* 'middle') and *geard* (yard nowadays) is cognate with Lat. *hortus* 'garden', from PIE **ǵher-* 'enclose'. *moncynnæs*: 'mankind', genit. sing., compound of *monn* 'man' < PIE **manu-* (cp. Ved. *mānu-* 'man') and *cynn* 'kin' < Gmc. **kunja-* < **ǵnh₂-jo-*, from **ǵenb₁-* 'give birth'. *Frēa*: 'Lord', Gmc. **frawōn*, the feminine of which became German *Frau* 'woman'.

Old High German

15.67. The language known as High German (loosely, German; it is called "High German" after the relatively mountainous terrain of much of its early territory, to distinguish it from the "Low German" of the lowland country to the north) appears

first in the form of runic inscriptions in **Old High German** dating back to 600 or so, most importantly on a spearhead found in **Wurmlingen near Tübingen** in southwest Germany. From the **eighth century** come a large number of glosses as well as the earliest literature, at first short poems and religious texts. But only in the ninth and tenth centuries does a considerable quantity of material appear on the scene, roughly contemporaneous with the flowering of Old English.

15.68. Old High German is preserved in six dialects, which often vary substantially from one another in orthography and certain grammatical features. The dialect called **East Frankish**, spoken in east-central Germany east of Frankfurt, is often treated by modern scholars as a kind of standard Old High German, but at the time there was no literary standard that obtained across different regions. The other dialects (Old Alemannic, Old Bavarian, South Rhine Franconian, Rhine Franconian, and West Frankish) were collectively spoken in an area stretching from Switzerland and Bavaria in the south up through Cologne and Fulda in the north.

Often considered a separate West Germanic language, but forming part of the general Old High German dialect area, was the **Frankish** spoken by the Franks, a people that progressively invaded areas of the Roman Empire, primarily eastern Gaul, beginning in the fourth century. The empire they formed, under the Merovingians and Carolingians, lasted more than 300 years, and in the west they (and their name) ultimately became part of the nationality now known as the *French*. The western variety of Frankish is not attested with certainty, but it exerted strong lexical influence on French.

Another variety of Old High German was the language of the Lombards, called **Langobardic**, about which we know very little. The Lombards came from what is now northwest Germany and migrated southward during the fifth century, eventually establishing an important kingdom in northern Italy.

15.69. The surviving corpus of Old High German literature was produced in monasteries and is almost exclusively Christian. We have only the meagerest remains of pagan Old High German literature, including the charm below in §15.76 and a solitary poem in traditional Germanic heroic verse, the incomplete, quasi-historical *Hildebrandslied*, which is intermixed with Old Saxon.

The High German Consonant Shift

15.70. The most famous hallmark of German that separates it from the other West Germanic languages to its north is the **High German Consonant Shift**, or as it is often called in German, the *zweite Lautverschiebung* (Second Sound Shift, that is, second after Grimm's Law). By this sound change, broadly speaking, the West Germanic voiceless stops *p *t *k were affricated to *pf ts* (written *z* or *tz*) and *kx* in initial position, and spirantized (became fricatives) to *f s x* (written *ch*) elsewhere. A comparison of words in modern English and German will illustrate these changes in word-initial (left) and word-internal (right) positions:

English	German	English	German
pound	Pfund	ape	Affe
tin	Zinn	water	Wasser
cook	Kchoch (Swiss German)	make	machen

As shown by the 'cook' entry, only Swiss German shifted **k* to *kx* in initial position; 'cook' in standard German is *Koch*. In fact, as one travels northward in Germany, ever fewer components of the High German Consonant Shift are met with, until one arrives in Low German territory, where none of the shift has taken place.

15.71. High German is also characterized by the shift of West Germanic **d* to *t* and **þ* to *d*: compare Eng. *deed* with German *Tat* and Eng. *thing* with German *Ding*. This change also affected Dutch and spread to Low German, although the latter's ancestor, Old Saxon, does not show either change.

Changes to vowels

15.72. Old High German was much more conservative in its vowel system than Old English. One change that may be noted is the diphthongization of **ō* to *uo* in Old High German, which became long *u* in the later language. This is the source of the difference in the vowels in such English–German pairs as *brother/Bruder*, *foot/Fuß*, *book/Buch*, and *flood/Flut* (OE *brōþor*, *fōt*, *bōc*, *flōd* vs. OHG *bruoder*, *fuoz*, *buoh*, *fluot*).

Middle and Modern High German

15.73. Around 1100 was the birth of the Middle High German period. For the first time, a standard German was developed, based on the Bavarian and Alemannic dialects in the south. The most important Middle High German literature is poetry, in particular courtly epic and romance, stressing chivalric virtues and refinement and exploring Christian themes. The courtly poetry also included the important genre of the *Minnesang* or song of courtly love (*minne*). Another genre, that of the national epic, is most famously represented by the anonymous *Nibelungenlied*, written in Austria c. 1200–1210; this work is rooted in the earlier Germanic warrior tradition and draws heavily upon historical and legendary material about the Danes, Saxons, Burgundians, and Huns in the fifth and sixth centuries.

MODERN HIGH GERMAN

15.74. For several hundred years after the height of Middle High German literature, there was no longer any standard literary language. By far the most important influence on the development of the Modern High German standard language was Martin Luther's translation of the Bible, the first edition of which appeared in 1522 (Old Testament) and 1534 (New Testament). Luther's translation was the first to be written in a direct and uncomplicated – at times even colloquial – style that strove not only to include expressions that were modern and up-and-coming, but also to incorporate linguistic features from as many regions as possible. Its impact on literary German was immense; its core was Luther's native dialect of Thuringian.

YIDDISH

15.75. During the Roman Empire, Jews had settled in various places in Europe, bringing the Semitic languages Hebrew and Aramaic with them from Palestine; wherever they settled they acquired the local language but continued to use Hebrew as a religious language. By the late Middle Ages, the Jews in Germany were primarily

settled in east-central and southern parts of the country, and their German was thus Bavarian or East Middle German, a variety spoken north of Bavaria. Their German contained Hebrew lexical elements mixed in with it; it is this mixture that is known as Yiddish ('Jewish' in its original literal meaning, German *Jüdisch*). Historically it is a form of German, not Hebrew; its sentence structure has always been German, as is about three-fourths of its vocabulary. After many of its speakers moved east into Slavic territory, it became influenced by Polish and other Slavic languages. The first documents in Yiddish date to the twelfth century; the language is usually written using the Hebrew alphabet.

Old High German text sample

15.76. The Second Merseburg Charm, from the tenth century. The two Merseburg Charms are the only surviving examples of pre-Christian pagan magic spells in German. The last two lines have numerous analogues in other IE traditions, including Anatolian, Indic, Celtic, and Tocharian. The story of a god's horse spraining its foot and being healed by Wotan is apparently depicted also on a golden bracteate (a kind of pendant or amulet) from c. AD 500.

Phol ende Uuōdan uuorun zi holza.
 Dū uuart demo Balderes uolon sīn uuoz birenkit.
 Thū biguol en Sinthgunt, Sunna era suister,
 thū biguol en Friia, Uolla era suister,
 thū biguol en Uuōdan sō hē uuola conda:
 sōse bēnrenkī, sōse bluotrenkī,
 sōse lidirenkī:
 bēn zi bēna, bluot zi bluoda,
 lid zi geliden, sōse gelimida sīn.

Phol and Wodan were riding to the woods,
 and the foot of Balder's foal was sprained.
 So Sinthgut, Sunna's sister, conjured it;
 and Frija, Volla's sister, conjured it;
 and Wodan conjured it, as well as he could:
 Like bone-sprain, so blood-sprain,
 so joint-sprain:
 Bone to bone, blood to blood,
 joint to joints; so may they be glued.

15.76a. Notes. Phol: an otherwise unknown name, but probably the masculine of *Uolla* below. The names Sinthgut and Sunna are also found only here. ende: 'and', PIE **h₂enti*, also in Gk. *anti* 'against', Lat. *ante* 'before'. Uuōdan: Wotan, the supreme Germanic god (Odin in Scandinavia). The English equivalent, *Woden*, is the source of *Wednes-day*. uuorun: 'rode', Mod. HG *fuhren*, from Gmc. **fōrun*, a Class VI strong verb (§15.26), from PIE **per-* 'cross over'; cognate with Eng. *fare*. zi: 'to', Mod. HG *zu*. holza: 'wood, forest', dat. sing. dū: 'then, there, so, and', often little more than a sentence-connector. uuart: 'became', early Mod. HG *ward*. demo: 'to the', dat. sing. of the definite article, modifying *uolon*. Balderes: 'of Balder' (or Baldur), son of Odin. uolon: 'foal', dat. sing. The construction is impersonal, literally 'to Balder's foal his foot was sprained'. sīn: 'his'. uuoz: 'foot', Mod. HG *Fuß*. birenkit: 'sprained', past participle, Mod. HG *gerenkt* (with different prefix). Cognate with Eng. *wrench*. biguol:

'conjured', preterite 3rd sing. **en**: 'him, it', accus.; Mod. HG *ihn*. Eng. *him* continues the dative case, OHG *im*. **era**: 'her', Mod. HG *ihre*. **suister**: 'sister', Mod. HG *Schwester*. Eng. *sister* is a borrowing from Norse; the OE form was *sweostor*. The *-t-* is epenthetic, arisen in early Germanic in the *st-*cluster of the inherited weak stem **syes-*. **Frīia**: 'Freyja', the goddess of love (ON *Frigg*). **Uolla**: perhaps the same as *Fulla*, a handmaid of the Norse goddess *Frigg*. **sō**: 'as, so'. **hē**: 'he', not the same form as Mod. HG *er*. **uola**: 'well'. **conda**: 'could, knew how', Mod. HG *konnte*. **sōse**: 'just as'. **bēnrenkī**: 'bone-sprain'. **bluotrenkī**: 'blood-sprain'. **lidirenkī**: 'joint-sprain' or 'limb-sprain'. Note that in this and the following part of the charm, the body-parts are enumerated going from inside (bone) to outside (joint or limb), via flesh (blood), the direction for driving out the ailment. The same order of body-parts is found in spells and medical texts elsewhere in IE, and in fixed traditional descriptions of the so-called "canonical creature," and is likely inherited. **geliden**: 'joints' or 'limbs'; instead of the usual plural we get a form called a collective, commonly formed in Germanic with the prefix **ga-* (> German *ge-*). The collective was eventually generalized in Mod. HG *Glied* 'joint, limb'. **gelimida sin**: 'they may be glued' or 'they are glued'. The verb is related to Mod. HG *Leim* 'lime, glue', and *sin* is a 3rd pl., either subjunctive (Mod. HG *seien*) or indicative (*sind*).

Old Saxon

15.77. Old Saxon was spoken until about the twelfth century in an area bounded by the Rhine in the west and the Elbe in the east, and from the North Sea down to Kassel and Merseburg in the south. Its most significant literary remains are the *Heliand* ('Savior'), composed about 830, which tells the life of Christ in almost 6,000 lines of alliterative Germanic verse. We also have part of an Old Saxon translation of the Book of Genesis, plus some minor fragments.

15.78. Linguistically, Old Saxon is quite close to Old English, Old Frisian, and Old Low Franconian (the ancestor of Dutch), being set apart from High German by the lack of the High German Consonant Shift (§15.70 above). It differs from the rest of West Germanic in preserving *ð* unchanged word-internally, as in *liobora* 'better, dearer' in the text sample below (contrast German *lieber*). Additionally, Old Saxon lost *h* in the cluster *hs*, as in the word for 'flax', *flas* (OE *fleax*, OHG *flahs*). Old Saxon does not show the effects of *i*-umlaut as thoroughly as Old English or Frisian, at least not in spelling; for example, the plural of *mūs* 'mouse' is also *mūs* (vs. OE *mýs*, from **mūsiz*).

15.79. The modern-day descendants of Old Saxon are the varieties of Low German spoken in northern Germany. They blend into Frisian and Dutch in the west and northwest. Low German reached its zenith as a literary and administrative language during the time of the Hanseatic League, a powerful commercial confederation of cities in northern Germany and along the Baltic coast whose height of power was from the thirteenth to the fifteenth centuries. The center of the Hanseatic League was Lübeck, whose variety of Low German (Eastfalian or Ostfälsch) became widely used as an administrative and literary language, and was for a time more important than Middle High German.

Old Saxon text sample

15.80. Excerpt from the story of Jesus's return to Galilee in the *Heliand* (1121–27 and 1148–50). The biblical stories contained in the *Heliand* are skillfully recast in a Germanic setting. The characters are given social roles proper to aristocratic

Germanic society: Jesus, for example, is the “giver of jewels,” and he is “chosen” by James and John as Lord in the same way leaders were elected in early Germanic society. Herod is similarly depicted as a “ring-giver” who holds feasts with his “ring-friends,” both central Germanic images of a king. The Lord’s Prayer is likened to runes or secret knowledge; in teaching it, Jesus calls to mind Woden, who knows the secrets of magical formulations.

Here Jesus is depicted like a Germanic chieftain assembling a retinue of young warriors.

- Uuas im an them sinuuelði sâlig barn godes
 lange huile, unthath im thô liobora uuarð,
 that he is craft mikil cûdien uolda,
 uueroda te uuillion. Thô forlêt he uualdes hlêo,
 1125 ênôdies ard endi sôhte im eft erlo gemang,
 mâri meginthioda endi manno drôm,
 geng im thô bi Iordanes staðe . . .
 1148 . . . He began im samnon thô
 gumono te iungoron, gôdoro manno,
 1150 uuordspâha uueros.

(1121) The blessed son of God was in the wilderness (1122) for a long time, until it then seemed better to him (1123) that he should make known his great strength (1124) for advantage to the people. Then he left the shelter of the forest, (1125) the domicile of the wilderness, and sought a group of earls again, (1126) famous great people and the tumult of men, (1127) (and) he went then to the bank of the (River) Jordan . . . (1148) He began to gather there (1149) men for disciples, good men, (1150) word-wise men.

15.80a. Notes (very selective). *uualdes hlêo*: ‘shelter of the forest’; *uuald* ‘forest’ is cognate with German *Wald*, and *hlêo* is cognate with Eng. *lee* (‘sheltered side’, OE *hlêo*). The desert of Judaea is transformed by the author of the *Heliand* into a more Germanic-like forested region. *gumono*, *manno*, *uueros*: three different words for ‘man’, all of them inherited from PIE (**dhǵhemō(n)*, **manu-*, **miH-ro-*). *uuordspâha*: ‘word-wise, eloquent’; word-wise in order to convert people.

Dutch and Frisian

15.81. To the west of Old Saxon was spoken Old Low Franconian, the speech of a group of western Franks and closely related to Old Saxon. It is attested, very meagerly, from the ninth to the twelfth centuries in Limburg in the extreme southeast of the modern Netherlands. The particular variety of Old Low Franconian spoken in the cultural centers of Flanders and Brabant is called Old West Low Franconian and is the ancestor ultimately of Dutch. It is known only from a charming two-sentence scrap preserved in the binding of an eleventh-century Latin manuscript in England: *hebban olla vogala nestags hagunnan hinase hi[c]e nda thu w[at u]nbidan [w]e nu* “All the birds have begun nests except for you and me. What are we waiting for?”

15.82. The medieval descendant of this language was Middle Dutch, which is well attested starting in the late twelfth century, principally in Flanders and Brabant.

The literature of this time consists of courtly romances and epics, miracle plays, and religious writings. In the seventeenth century the Dutch colonized South Africa; their descendants speak Afrikaans, which has diverged from Dutch significantly enough to be considered a separate language. The name of the dialect that developed around Flanders, namely Flemish, refers to the Dutch spoken in Belgium. Dutch, Afrikaans, and Flemish are often together called *Netherlandic*.

15.83. In northern coastal areas of the Netherlands and Germany, as well as on the North Sea islands off the coast, is spoken Frisian. Frisian is said to be the closest living relative of English (recall §15.52). Its first records, in Old Frisian, date only to the thirteenth century, although some scattered runic remains from the sixth through the ninth centuries have been claimed (doubtfully, at least in part) to be in an earlier form of the language.

North Germanic: Old Norse and Scandinavian

15.84. North Germanic is represented by a single ancient language, Old Norse, and its descendants the modern Scandinavian languages. It is probable that Runic, or Runic Norse (see §§15.33ff.), was its immediate ancestor. The homeland of the North Germanic speakers was centered in an area along the western Scandinavian coasts, especially Norway and southwest Sweden but also northern Denmark. The northern Germanic pirates known as Vikings mostly spoke varieties of Old Norse; for reasons that are still unclear, around the seventh century these Norsemen began a series of raids that soon grew into a scourge as they ravaged and terrified any part of Europe that was reachable by boat. They had developed the best nautical technology then known and perfected the technique of the lightning raid; coastal populations were essentially defenseless against them. Raiding per se is an old Indo-European tradition; there is evidence suggesting that it was a rite of passage for young unmarried warriors to enter into a *Männerbund*, a 'band of men' or fraternity of sorts, and undertake marauding expeditions (§2.8). From the seventh through the tenth centuries and beyond, Viking bands ventured as far afield as Paris, Pisa, Newfoundland, Jerusalem, and even Baghdad. But the Norsemen were often more interested in trading than raiding, notably those that went east and settled in what is now Russia, founding such cities as Novgorod and Kiev and becoming leaders of Rus', the medieval Russian state.

The Vikings that laid siege to Paris withdrew upon being granted a duchy in northern France, now called Normandy (from *Norman*, a variant of *Norseman*) in 911. The Norsemen that settled there quickly took on French ways and customs, but the restless adventurousness in their blood remained: their descendants were the Normans who seized the throne of England in 1066.

15.85. In England, the Norsemen settled along large stretches of the north and central coast, an area known as the Danelaw. The names of the cities and towns of this area are mostly of Norse origin and contain such suffixes as *-by* ('town, village', from ON *býr*, whence also Eng. *by-law*) and *-thorpe* ('village', ON *þorp*). Norsemen and Anglo-Saxons intermarried in this area; it is sociolinguistically significant that the word *husband* is Norse, while *wife* is Anglo-Saxon – a fact attesting to the widespread taking of Anglo-Saxon wives by Norse settlers. Hundreds of Norse words entered English, slowly filtering through to southern England during the Middle English period, and including such basic vocabulary items as *get*, *sister*,

take, both, call, and they. The influence of Norse on English core vocabulary was actually greater than that of Norman French.

Old Norse literature

15.86. The Runic language of the earliest runic inscriptions, as discussed above, may be ancestral to Old Norse (see §15.34). The first inscriptions that are unambiguously Norse begin to appear around the seventh century; not much later clear dialectal differences within Norse can be detected, allowing us to distinguish among Old Icelandic, Old Norwegian, and Old Swedish.

Old Norse literature occupies a special place in the hearts of Germanic philologists because it preserves native pre-Christian Germanic mythology and folklore far better than any other old Germanic literature. Technically, most Old Norse literature is written in the dialect spoken in Iceland and called Old Icelandic; often the term "Old Norse" refers really to this dialect. Written literature in it dates from the mid-twelfth century, and consists of historical accounts of the travels and raids of the Norsemen, as well as mythological poetry and prose, some of it composed orally several centuries earlier. The most important of the poetry is the poetic (or Elder) Edda, a collection of mythological poems that probably represent our earliest Old Norse literature. The historian and chieftain Snorri Sturluson (1179–1241) is the source of the equally famous Prose Edda, containing prose versions of many other mythological stories compiled as part of an effort to preserve these tales for posterity. Old Norse prose is also represented famously by the sagas (*saga* 'something said, tale'), written between the twelfth and fifteenth centuries; these represent one of the highpoints of medieval European vernacular literature. Finally, mention may be made of the poetry of the Skalds (ON *skáld* 'poet'), court poets in Norway who fashioned several new poetic genres and styles; skaldic poetry is known especially for its dense and elaborate use of kennings or telescoped metaphors, which can make it quite challenging to read (e.g. *lyngfisk* 'heather fish' = 'snake', crawling in heather like a fish in the sea, or *Haka bláland* 'blue land of Haki' = 'sea' [Haki is the name of a sea king]). Kennings are in fact a characteristic feature of traditional Germanic poetry (as e.g. OE *gūþwudu* 'battle-wood' = 'shield'), but nowhere is the art of the kenning so highly developed as in skaldic poetry.

Developments of Old Norse

15.87. Old Norse looks somewhat different from the other Germanic languages of comparable age; impressionistically, the words seem shorter and the vowel inventory greater. This is due to several phonological changes that affected the language between the time of the early runes and the eighth or ninth century.

VOWELS

15.88. Syncope. The vowels of many final syllables were syncopated (deleted), as in ON *gestr* 'guest' from **gestir* (Runic *gastiz*), and the nominative singular ending -*az*, which already in later Runic inscriptions appears as just -*z*. Before this syncope had run its course, however, the language underwent a series of vowel mutations or umlauts, described next.

15.89. *u*-Umlaut. Certain vowels were rounded preceding a syllable with a *u* in it, a change called *u*-umlaut or *u*-mutation. An *a* was rounded to a sound written *ø*; *e* was rounded to *ø* (pronounced like German *ö*); and long *á* was changed to long *ô*. (Acute accents indicate length in standard Old Norse orthography.) So for example North Germanic **saku* became Old Norse *søk* 'charge, offense' (cp. OE *sacu*, Mod. Eng. *sake*); similarly, the verb *róa* 'to row' has a reduplicated 3rd pl. preterite *røru* 'they rowed' from **rerun*.

15.90. *w*-Umlaut. Closely allied with *u*-umlaut was *w*-umlaut, caused by a following *w* (later becoming *v*). In addition to the above effects, this also changed *i* to *y* (pronounced like German *ü*), as in the verb *slyngva* 'to fling, sling'.

15.91. *i*-Umlaut. Norse had *i*-umlaut just like its West Germanic cousins, whereby an *i* or *j* (the glide *j*) fronted a vowel in a preceding syllable. The effects of this can already be seen in later Runic. The umlauting of the back vowels *a*, *o*, and *u* and their long counterparts proceeded as in West Germanic (§15.48), with the outcomes spelled *e* *ø* *y* for the short vowels and *æ* *œ* *ý* for the long ones. Additionally, the diphthongs **au* and **eu* (which developed into *jú*) were umlauted to *ey* and *ý*, and the vowel *ø*, a product of *u*-umlaut, was fronted to *ø* like *o*. Examples include *framr* 'forward' ~ *fremri* 'more forward'; *áss* 'god' (nomin. sing.) ~ *æsir* 'gods' (nomin. pl.); *koma* 'to come' ~ *kømr* 'comes' (**komiz*); *fullr* 'full' ~ *fylla* 'to fill' (**fulljan*); *fljúga* 'to fly' ~ *flýgr* 'flies' (**fleugiþ*).

15.92. Breaking of *e*. Finally, we may mention the "breaking" of *e*, which was analogous but not identical to the breaking seen in Old English: *e* became *ea*, then *ja*, in a syllable followed by *a*, as in *jafn* 'even' from **ebna-* (cp. OE *efen*); and it became *eo* (later *jo* and even later *jø*) before a syllable containing a *u* (e.g. *jøtunn* 'giant' < **etunaz*).

CONSONANTS

15.93. For the consonants, a few characteristic changes may be mentioned. The most problematic for issues of Germanic subgrouping has been the "hardening" of the geminated glides **ww* and **jj* (see above, §15.41). The old *z* of Runic became *r* as in West Germanic, e.g. *vōro* 'they were' < **wāzun*. Word-initial *j* disappeared, as in *ár* 'year' (cp. German *Jahr*) and *ungr* 'young' (cp. German *jung*), which happened also to *w* word-initially before rounded vowels and resonants, as in *orð* 'word', *urðu* 'they became' (cp. German *wurden*), and *lita* 'face' (cp. OE *wlitan* 'face'). Where *w* was not lost, it became *v*, as in ON *hvat* 'what', *vatn* 'water'.

15.94. Assimilation of consonant clusters. A major source of the unusual look of Old Norse vis-à-vis the other old Germanic languages is the great number of consonant clusters that underwent assimilation. The rules are too complex to enter into here; some examples will suffice: *sótti* 'sought' (cp. OE *sōhte*); *finna* 'to find' (cp. OE *findan*); *batt* 'bound' (cp. OE *band*); *villr* 'wild' (cp. OE *wild*); *Pórr* 'Thor' (literally 'thunder' < **Þunraz*, cp. OE *Þunor*); and *drekka* 'to drink' (cp. OE *drincan*).

Morphology

15.95. Old Norse nouns and verbs show the same behavior and inflectional categories as in contemporary relatives like Old English and Old High German. Due to the massive series of umlauts, the vowel alternations in stems can be complex, even more so

than in Old English (see §15.61): for example, the nominative, dative, and genitive singular of the word for 'father' was *faðir*, *feðr*, *foður*. Verbs do have one significant inflectional difference from the rest of Germanic in that the 3rd singular ending in the present tense was *-z, displacing the inherited ending *-þ. This spread from the 2nd singular. Hence *berr* (< *ber-iz) means both 'you carry' and '(s)he carries'.

ENCLITIC ELEMENTS

15.96. Characteristic of Old Norse is a high degree of enclisis of pronouns and adverbs that in the other Germanic languages remained separate words. The Norse definite article is ordinarily suffixed to the word it modifies: masc. -(i)n (e.g. *úlfr-inn* 'the wolf'), fem. -(i)n (e.g. *kona-n* 'the woman'), neut. -(i)t (e.g. *auga-t* 'the eye'). Old Norse created a class of reflexive or middle verbs by suffixing -mk (1st person) and -sk (2nd and 3rd persons), reduced forms of the reflexive pronouns *mik* and *sik*, as in *kalla-sk* 'calls himself/herself, is called' and *kollu-mk* 'I am called'.

15.97. Of some historical interest are the numerous Norse negators. Originally they were not negators at all, but were used to emphasize negative words; later they acquired negative meaning of their own. A negative suffix -gi is found in *engi* 'no, none' (*ein* 'one' + -gi), *ekki* 'nothing' (neut. *eitt* of *ein* + -gi), *eigi* 'not' (*ei* 'ever' + -gi), and *hvergi* 'nowhere'. The suffix -gi comes ultimately from PIE *k^wid, in its use as an indefinitizing suffix (e.g. Latin *quis-quis* 'anyone'; recall also the history of Gk. *ou* 'not' in §7.24). Of particular interest are the enclitic negative suffixes -a and -at attached to verbs, sometimes together with enclitic suffixed subject particles, e.g. *es-a* 'is not', *vas-k-a* 'I was not' (-k, enclitic form of *ek* 'I'), *hef-k-at-ek* 'I have not' (with doubled enclitic pronoun). Etymologically these are from unstressed *ein* and *eit* 'one'. The acquisition of negative meaning by words that were not negative to begin with is paralleled in numerous other languages, as with French *pas* (recall again §7.24).

The modern Scandinavian languages

15.98. Beginning about the ninth century the Runic inscriptions in Scandinavia begin to show regional differences, especially a general split into a West Nordic and an East Nordic variety, corresponding to the modern groups of West Scandinavian and East Scandinavian. Overall, the western dialects are more conservative than the eastern ones.

15.99. In the ninth century, explorers from the western Norwegian coast settled Iceland; their language developed into what is now known as Icelandic. The language of Iceland stayed remarkably conservative; speakers of modern Icelandic can read the Old Norse sagas and other literature without great difficulty. Modern Icelandic still retains three genders, four cases in the noun, and essentially the same pronominal and verbal system as in Old Icelandic. In pronunciation, however, there have been extensive changes. Close to Icelandic is Faroese, the language of the Faroe Islands between Iceland and the Shetland Islands, settled also by western Norwegians around the same time as Iceland. A handful of Faroese runic inscriptions are known from c. 1000–1500, as well as a few documents from the end of the thirteenth century; overall, though, little Faroese literature was produced until the nineteenth century, when Danish ceased to be the standard written language.

15.100. The North Germanic dialect of Norway developed into Norwegian, first recognizable as such (in the form of Old Norwegian) in the twelfth century.

Norwegian is notable for having developed contrasting tonal contours in words of more than one syllable, called accent 1 (falling-rising; developed on words that were originally monosyllables like *kommer* 'comes' < ON *kómr*) and accent 2 (rising-falling-rising; developed on words that were originally polysyllabic, as in *gater* 'streets' < ON *gǫtur*).

15.101. To East Scandinavian belongs first of all Danish, arguably the least conservative Scandinavian language, which can be inscriptionally distinguished by 1000 or so (Old Danish), with the first Danish manuscripts dating to the thirteenth century. During the Middle Ages, it merged the Old Norse masculine and feminine genders into a single (common) gender, lost all the cases besides the nominative and genitive, and came under considerable influence from the neighboring Low German dialects to the south, from which it has made extensive morphological borrowings. Most Danish dialects have an unusual phonetic feature called *stød*, which is similar to a glottal stop near final consonants in monosyllables.

15.102. Around the same time as Old Danish came into existence we can also distinguish Old Swedish. Swedish, like Danish, lost most of the cases in the noun, but most dialects except the standard language have kept the three-way gender distinction. It also has the accentual contrasts of Norwegian.

Old Norse text sample

15.103. The opening of the *Prymskvíða* or *The Lay of Thrym*, a famous tale in the Poetic Edda (given here in the linguistic form of its probable tenth-century origin rather than that of the manuscript from 1270).

- 1 Vreiðr vas þá Ving-Þórr es vaknaði
ok síns hamars of saknaði;
skegg nam at hrista, skǫr nam at dýja,
reð Jarðar burr um at þreifask.
- 2 Ok hann þat orða alls fyrst of kvað:
'Heyrðu nú, Loki, hvað nú mælik,
es engi veit jarðar hvergi
né upphimins: Áss es stolinn hamri!'
- 1 Enraged was Ving-Thor, when he woke up
and missed his hammer;
he began to shake his beard, he began to shake his hair,
the son of Earth groped about himself.
- 2 And he said first of all these words:
'Hear now, Loki, what I now say,
which no one knows anywhere on earth
nor in heaven above: the god has had his hammer stolen!'

15.103a. Notes. 1. *vreiðr*: 'enraged', cp. Eng. *wrath*. The manuscript has just *reiðr*, showing the loss of *v-* before liquids; but when this poem was composed, the *v-* must have still been there because the word has to alliterate with *Ving-*. *Ving-Þórr*: 'Swinging Thor', referring to Thor's attribute of swinging his hammer. *waknaði*: 'woke up, awakened', intransitive. ON

has a class of verbs in *-na-* that have intransitive inchoative meaning ('become X'); exactly how they connect with other nasal presents in Germanic and IE is not fully clear. *ok*: 'and', cp. German *auch* 'also'. *hamars*: 'hammer', genitive object of the verb *sakna* 'to miss'. *nam*: lit. 'took', here more weakly 'did'; cognate with German *nahm* and used in Eddic poetry as an auxiliary for metrical convenience. *at*: 'to' with infinitives. *rēð*: lit. 'counseled', also 'set about (to...)' or just 'did', so here 'set about to grope' or just 'did grope'. The infinitive is *ráða*, cognate with German *raten* 'counsel, advise'. *Jarðar*: 'Earth', mother of Thor; nomin. sing. *jorð*. *þreifask*: 'to grope'; *-a* is the infinitive ending, cognate with German *-en* (OHG *-an*), and *-sk* is a reflexive suffix; §15.96.

2. *orða*: 'words', genit. pl. with *þat* (lit. 'that of words'); see §15.93. *kvað*: 'spoke', cognate with Eng. *quoth*. *heyðu*: 'Hear you', a combination of *heyr* 'hear' and the 2nd sing. pronoun *þú*. *Loki*: Loki, a god and trickster figure. *mælik*: 'I say', combination of *mæli* and suffixed clitic pronoun *-k* (cp. Runic *-ka*, the suffixed form of *ek*). *engi*: 'no one', originally 'anyone', but over time it achieved negative force by being always used with a negative, like French *personne* 'no one', lit. 'person'. *veit*: 'knows', German *weiß*, Goth. *wait*. *hvergi*: 'anywhere, nowhere', composed of *hver* '(some)where' and the negative particle *-gi*; see §15.97 above. *upphimins*: 'up-heaven, the sky above'. The word is part of a traditional formula meaning 'earth and up-heaven' that also appears in OE and OHG. *Áss*: 'god', nomin. sing., stem *ás-*; pl. *Æsir*. The word comes from PIE **h₂ens-*, having derivatives pertaining to royalty and deities, such as Hitt. *ḫaššus* 'king' and Ved. *āsu-ra-* 'lord'. The last sentence literally means 'The god is robbed as to (his) hammer'.

For Further Reading

The debate about the internal subgrouping of the Germanic branch is thoughtfully treated in Nielsen 1989 and 2000. For beginners coming at Germanic from the English point of view, see especially Robinson 1992. A good overall history of Germanic, though dated in places, is Prokosch 1939; more recently, see the linguistic and cultural overview in Green 1998. Bammesberger 1990 is a very useful overview of Germanic nominal morphology, with many comparative lists; similar but not as good is Bammesberger 1986, on the verb. An excellent exhaustive comparative listing of the Germanic strong verbs is Seebold 1970.

The standard edition of the older Runic inscriptions known by the mid-1960s is Krause and Jankuhn 1966 (see also Krause 1971 for grammar); a bit more recent, and in English, is Antonsen 1975, although some of his interpretations must be treated with caution. Antonsen's book includes a brief grammar of Runic. The complete Gothic Bible, with glossary, is in Streitberg 2000; of several good available comparative grammars of Gothic, Krause 1968 and Braune 1981 may be mentioned. The standard grammar of Old High German is Braune 1987, with an accompanying reader (Braune 1994). For Old English, consult the classic grammar of Alistair Campbell (Campbell 1959); and for Old Saxon, see Holthausen 1921. The standard introduction to Old Norse in English, with a wide selection of texts and glossary, is Gordon 1962, also with a (rather bare) grammatical sketch. In German, see also the grammar of Noreen 1923.

Many of the standard etymological dictionaries are in German. For Old English, Holthausen 1934 lists cognates in and outside Germanic but few reconstructions; for the modern language, see Watkins 2000 (discussed in chapter 2). A Gothic etymological dictionary is Lehmann 1986, although the IE reconstructions deviate somewhat from standard notation. An up-to-date and exhaustively researched Old High German dictionary, with full etymological information and contributions by several major Indo-Europeanists, is underway and has progressed through the letter E: *Althochdeutsches Wörterbuch* (Berlin: Vandenhoeck und Ruprecht, 1952–). Quite useful for German and Germanic etymology is Kluge 2002. Old Norse etymology is well served by de Vries 1962.

For Review

Know the meanings or significance of the following:

Northwest Germanic	weak verb	Wulfila	High German
Grimm's Law	Verner's variant	Verschärfung	Consonant Shift
Verner's Law	dental preterite	i-umlaut	<i>Heliand</i>
Rasmus Rask	preterite-present	West Saxon	Danelaw
strong adjective	"V2"	<i>Beowulf</i>	Eddas
weak adjective	rune	breaking	u-umlaut
strong verb	futhark	Great English	
	Crimean Gothic	Vowel Shift	

Exercises

- 1 Give the Germanic outcomes for the following PIE sounds. Some may have more than one answer.

a	*b	c	*o	e	*p	g	*ŋ	i	*dh	k	*k̑	m	*s	o	*gh
b	*ō	d	*t	f	*r	h	*kʷ	j	*y	l	*ā	n	*ei	p	*gʷh

- 2 Give the PIE sound from which the boldfaced sounds in the English words below are likely to have descended:

a	<i>leech</i>	c	<i>seep</i>	e	<i>stare</i>	g	<i>hollow</i>
b	<i>quoth</i>	d	<i>have</i>	f	<i>root</i>	h	<i>bloom</i>

- 3 Below are some slightly simplified stems in pre-Germanic and their outcomes in Modern English. (By "pre-Germanic" is meant here essentially PIE after laryngeal loss and compensatory lengthening, and the centum merger of velars.) Using §§15.6–9 and 15.15, apply the sound changes of Grimm's Law, Verner's Law, and the accent shift to each of the pre-Germanic forms. Do this step by step as in the following example:

Example: pre-Germanic **upélo-*, Modern English *evil*

answer:	<i>*upélo-</i>	→	<i>*ufélo-</i>	by Grimm I
	<i>*ufélo-</i>	→	<i>*ubélo-</i>	by Verner
	<i>*ubélo-</i>	→	<i>*úbelo-</i>	by accent shift

Not all the forms undergo change. Once a pre-Germanic sound has shifted according to one part of Grimm's Law, it will not be affected by any other part of Grimm's Law that happened earlier. A diphthong counts as a single vowel (or syllable) for the purposes of this problem.

In group (a), the English gloss is at the same time the English descendant form.

pre-Germanic	English	pre-Germanic	English
a			
*ábelo-	apple	*dhó-	do
*átalo-	Ethel(-red)	*ghé-	go
*tréi-	three	*gnó-	know
*tritió-	third	*gnéyo-	knee
*réise-	rise	*léuse-	lose
*roisjé-	rear	*dhrénge-	drink
*séðjo-	seat	*dhronglé-	drench
*sodjé-	set	*sténge-	stink
*loitjé-	lead	*stóngio-	stench
b			
*péku-	payment	*séute-	boil
*bhére-	carry	*sutonó-	boiled, soaked
*bhrronó-	(having been) carried	*lusonó-	lost
*upér-	above	*altó-	aged
*kasó-	rabbit		

4 For each of the forms in 3b above, try to figure out what the Modern English descendant is.

5 Given the changes that affected vowels in the history of Germanic and Old English as detailed in §§15.14 and 15.57, determine which ablaut grade in IE is continued by the Old English words given below. The Modern English descendant is given when different from Old English, but do not use it to answer this question. Some of the PIE forms have been slightly simplified. Aside from the sound changes given in the text, you will need to know the following: Gmc. *u > OE o except before r or m; and, for the purposes of this exercise, Gmc. *e > OE i. Remember too that PIE *ē > West Germanic *ā.

Example: Given the root *sed- 'sit' and OE *sadol* 'saddle', the *a* in OE must come from PIE *o, hence *sadol* continues a PIE o-grade (*sod-).

- a *leg- 'to collect, gather; say': OE *lēce* 'physician' (< *'one who speaks magic words') (early Mod. Eng. *leech*)
- b *lendh- 'land': OE *land*
- c *yes- 'to put on clothes': OE *werian* (Mod. Eng. *wear*)
- d *dhreibh- 'to drive, push': OE *drift*
- e *reġ- 'to direct': OE *riht* 'right' (Mod. Eng. *right*)
- f *reġ- 'to direct': OE *gerecenian* 'to arrange in order, account' (Mod. Eng. *reckon*) (Hint: in pre-OE the word would have been **gerecinian*)
- g *dhers- 'to dare': OE 3rd sing. *durst*
- h *reudh- 'to clear land': OE *rodd* 'stick' (Mod. Eng. *rod*)
- i *nem- 'to take': OE *numen* 'seized' (Mod. Eng. *numb*)

6 Provide a historical explanation for the boldfaced consonant alternations seen in the Modern English pairs *freeze* ~ *frore* (archaic, 'cold, frozen') and *seethe* ~ *sodden*.

- 7 The pre-Old English forms **stankja* and **drankjan* became Modern English *stench* and *drench*. Determine from this what chronological order the three sound changes of umlaut, loss of *j*, and palatalization happened in with respect to each other. More than one order is possible.
- 8 The form **drankjan* above literally meant 'to make drink'. What is its history, in Indo-European terms?
- 9 Account for the loss of the *n* in the OE 3rd pl. ending *-aþ* (§15.61), assuming the immediately preceding stage was **-āþ*.
- 10 Provide an explanation based in Indo-European morphology and Germanic phonology for why the preterite presents, uniquely among Germanic verbs in the indicative mood, lack an inflectional ending in the 3rd person singular.
- 11 Based on the Germanic word for 'daughter', **duhter-*, did Grimm's Law precede the loss of laryngeals, or vice versa? Explain your answer.

PIE Vocabulary VII: Position and Motion

- *h₁es-* 'be': Hitt. *ēšzi* 'is', Ved. *ásti* 'is', Lat. *est* 'is', Eng. *is*
**sed-* 'sit': Ved. *ásadat* 'sat', Gk. *hézomai* 'I sit', Lat. *sedeō* 'I sit', OCS *sěděti* 'to sit'
**legh-* 'lie': Gk. *lékhetai* 'lies', OCS *ležati* 'to lie'
**kel-* 'lie': Luv. *zlyari* 'lies', Ved. *śáye* 'lies', Gk. *keítai* 'lies'
**steh₂-* 'stand': Ved. *tíṣṭhati* 'stands', Gk. (Doric) *hístāmi* 'I stand', Lat. *stāre* 'to stand'
**h₁ei-* 'go': Hitt. *ī* 'go!', Ved. *éti* 'goes', Gk. *eími* 'I (will) go', Lat. *eō* 'I go'
**g^wem-* 'come, go': Gk. *baínō* 'I come', Lat. *ueniō* 'I come'
**sek^w-* 'follow': Ved. *sácate* 'follows', Gk. *hépetai* 'follows', Lat. *sequitur* 'follows'
**uegh-* 'convey': Hieroglyphic Luv. *wa-zl/a-* 'drive', Ved. *váhati* 'drives', Lat. *uehō* 'I convey', OCS *vezp* 'I drive'
**bher-* 'bear': Ved. *bhárati* 'carries', Gk. *phérō* 'I carry', Lat. *ferō* 'I carry', OCS *berp* 'I take'
**h₂eg^w-* 'drive, draw': Ved. *ájati* 'drives', Gk. *ágō* 'I lead', Lat. *agō* 'I drive, do', Arm. *acem* 'I lead', Toch. AB *āk-* 'go, lead'

Introduction

16.1. Little is known about the origin of the Armenians. The region in north-east Turkey that they have occupied throughout their known history is called *Ḫayaša* in Anatolian cuneiform texts from the mid-second millennium BC, a name that has been plausibly compared with *Hay*, the Armenian word for '(an) Armenian'. But from the available evidence it does not appear that the Armenians or their ancestors had settled this region by that early date. The Greek historian Herodotus and other ancient sources agree that they were newcomers into Anatolia; Herodotus identifies them with the Phrygians. While this identification is probably not correct per se, it is likely that they and the Phrygians belonged to the same migratory waves of Balkan immigrants that started coming into Anatolia in the late second millennium BC. The handful of Anatolian loanwords in Armenian, such as *xalam* 'skull' (cp. Hitt. *ḫalanta-* 'head'), were probably picked up during the migration eastward through Anatolia.

16.2. The region in which they settled, around Lake Van, was the home of the powerful state of Urartu in the early first millennium BC. (Its name survives, via Hebrew, as the name of Mount *Ararat*.) The Urartians were a non-Indo-European people related to the Hurrians (see §10.21); both Hurrian and Urartian furnished another set of loanwords into Armenian, such as *xn̄jor* 'apple' (Hurrian *ḫinzuri*). The Urartian kingdom fell to Assyria in the eighth century BC, although its culture still flourished until the Armenians took over the area in the seventh century BC. But Armenian hegemony did not last long; the region, together with all the other former Assyrian principalities, were soon swallowed up by the Medes, an Iranian people to the east. Armenia became one of the vassal states of the large Iranian confederacy that formed the basis for the Persian Empire of the Achaemenids (see §11.26). It is during the Achaemenid dynasty that Armenia finally enters history as a Persian province called *Armīniya-* in the Bisotun inscription of Darius the Great.

16.3. The cultural and political domination of Armenia by Iranian civilizations continued into Middle Iranian times. The linguistic impact of this was enormous. There was a prodigious infusion of Iranian loanwords, especially from Parthian, the language of the Arsacid dynasty (247 BC–AD 224; §11.38), as well as a sizeable number from Aramaic, an important administrative language both in the Persian Empire and in Middle Iranian times. The advent of Hellenism brought many Greek words into Armenian too, but it is really the Iranian loans that typify the lexical makeup of the classical language. They are considerably more numerous than the

Norman French loanwords in English, and misled two or three generations of Indo-Europeanists into thinking that Armenian was an Iranian language.

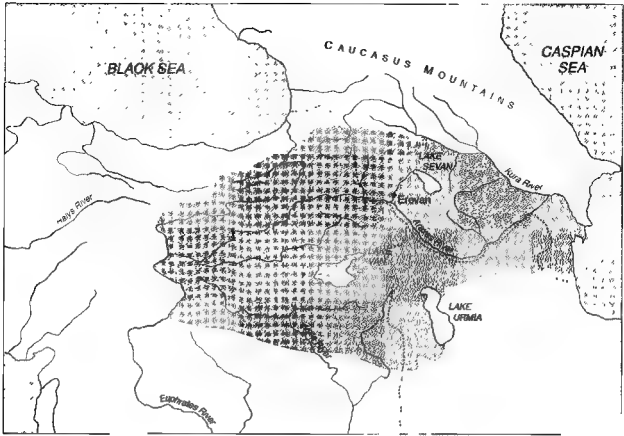
16.4. Not until 1877 was it proved that Armenian belonged to its own branch of Indo-European. The proof was furnished by a young scholar named Heinrich Hübschmann, then near the beginning of his career. His reasoning bears retelling, for it is an instructive example of good common sense. First, he pointed out that Armenian often possessed two words or morphemes for a particular concept, such as *jern*, the ordinary word for 'hand', alongside *dast*, which is borrowed from Persian and appears in a number of compounds. The Iranian words were easy to recognize, but once they were weeded out the residue of words that was left had systematically different outcomes of the PIE sounds from the Iranian outcomes. These words were then recognized by Hübschmann as the native lexical core of the language. Hübschmann further showed that the Armenian inflectional endings were wholly different from those of Iranian, and since inflectional morphology is rarely borrowed, these endings must also have been inherited.

In spite of the fundamental advances made in Armenian historical linguistics by Hübschmann and later scholars like Antoine Meillet, Armenian is still difficult for IE studies. This is primarily due to the small number of native forms left in the language by the time of its earliest attestation: no more than about 450 words are inherited. The small stock of native words has left precious few examples of many Armenian sound changes, some of which are among the most bizarre in the whole IE family (see below).

16.5. The relationship of Armenian to the other branches of IE has been much discussed. It shares certain features with Greek, Indo-Iranian, and Phrygian, leading numerous researchers to group these branches together as common descendants of a single dialect area of late PIE. But by the time we reach our earliest Armenian records in the fifth century AD (nearly two millennia after the earliest recorded Greek and Indo-Iranian), the evidence of any such early kinship has been reduced to a few tantalizing pieces.

16.6. The writing down of Armenian was a product of Christianization. Armenia was the first country to adopt Christianity as the state religion, around the year 300, replacing the Zoroastrianism that had been taken over from the Iranians. In the early fifth century, according to tradition, a cleric and scholar named Mesrob Maštoc' (or Mesrop Mashtots, 360?–440) devised an alphabet for writing Christian works in Armenian. The alphabet is largely based on Greek, and contains 36 letters (two additional ones were added in the twelfth century). It is excellently designed, as every sound in the language is represented by a single letter and vice versa, with only a very few exceptions. Mesrob used the alphabet to record the earliest translations into Armenian from Greek and Syriac; the Armenian translation of the Bible, which was partly due to Mesrob himself, was completed in the first half of the fifth century.

16.7. The fifth century is regarded as the golden age of Armenian literature and of the literary language called Classical Armenian or Grabar, which remained the standard into the nineteenth century. The best Classical Armenian is held to be that of the historian Eznik, whose *Against the Sects* (*Elc atandoc'*), written shortly before 450, is a valuable source of information on pre-Christian Armenian religion. Also from around this time are various histories, especially those of Elišē, P'awstos, Łazar of P'arpi, an anonymous historian known as Agat'angelos (Agathangelos), and Koriwn, who wrote a biography of Mesrob. Also very prominent is Movsēs Xorenac'i (Moses of Choren), who is important for Indo-European studies in preserving some



Map 16.1 Greater Armenia during the early Christian era

native folklore and poetry. Although Movsēs is traditionally dated to the fifth century as well, Western scholars now think he wrote considerably later, around the ninth century. Much of early Classical Armenian literature was translated from Greek or Syriac, and its often excessive faithfulness to the syntax of those originals even led to violation of the grammatical rules of Armenian itself. Our knowledge of Classical Armenian syntax is therefore not always on a solid footing. Aside from the few tantalizing bits preserved in Movsēs, we unfortunately possess essentially no pre-Christian Armenian literature.

16.8. No manuscripts from the early centuries of Armenian literacy survive, although there are inscriptions beginning already in the fifth century. The oldest known manuscript is the so-called Moscow Gospel, which was copied in 887, but most manuscripts are from a good deal later. (Eznik, for example, is only preserved in one manuscript that dates to 1280, over 800 years after he wrote.) As is always true in such cases, many modernizations and interpolations of later material have crept in.

16.9. The classical period was followed by the Middle Armenian period. The best-known variety of Middle Armenian is the official court language of the Cilician kingdom (1080–1375) in southeast Turkey. It is intermediate between the classical language and modern Western Armenian, one of the two modern literary languages. Western Armenian is based on the dialect around Istanbul; Eastern Armenian, the official language of Armenia, is based on the dialect around Mount Ararat and Erevan, the capital. The two literary varieties are fairly similar, but aside from these there exist scores of tremendously varied dialects, some of which have undergone significant phonological and syntactic transformation under the influence of neighboring Turkic and Caucasian languages. (The diaspora during and after the

genocide of Turkish Armenians in 1915–23 has brought speakers of most of the Western dialects to the United States.) The modern standard language has not been free of these influences either; in many areas of syntax, such as subordinate clausal structure, it more greatly resembles a Turkic language than an Indo-European one. Morphologically, modern Armenian has maintained essentially the same case system in the noun as the classical language, but has markedly changed verbal conjugation by introducing many periphrastic verbal constructions especially in the present and future tenses.

16.10. Classical Armenian is quite homogeneous across different authors from different areas, and may have been designed as a kind of regional standard. Supporting this hypothesis are indications of dialect mixture: the outcomes of several PIE sounds are not consistent, such as **p* and **s* (see §§16.13 and 23 below). The language had three parallel series of stops and affricates: a plain voiceless series *p t k c č* (the *c* represents the *ts*-sound of *bets*, and *č* is the *ch*-sound of *cheese*); a parallel voiced series *b d g j ĵ*; and a voiceless aspirated series transliterated as *p' t' k' c' č'*.

In addition to these consonants, Armenian had the fricatives *s z v* and *x* (the *ch*-sound of *Bach* in German), the nasals *m* and *n*, the glides *y* and *w*, and four liquids written *l r l̥ r̥*. The two *l*'s designate "light" and "dark" (velarized) *l*, the latter becoming later a voiced uvular fricative [ʁ] (a sound similar to French *r*). The *r̥* represented a trilled *r*, whereas ordinary *r* represented just a single tap.

16.11. There were seven vowels, transliterated as *a e ē i o ow ə*. The sequence *ow* represented the sound *u* (the Armenian convention of writing it as *ow* is borrowed from Greek, which used *ou* in its alphabet to write the same sound). Many authors transliterate *ow* as *u*, but we shall use *ow*. The schwa (*ə*) was only written at the beginning of a word (e.g. *ənd* 'with, along'), where it comes historically from a full vowel; but in the spoken language it frequently occurred preceding or breaking up consonant clusters, in which case it was not written (so *dpanem* 'I write' was pronounced *dəpanem*, and *xmnanē* 'from him' was pronounced *əxmnanē*).

From PIE to Classical Armenian

Phonology

CONSONANTS

16.12. Stops. Armenian is a satem language. The PIE labiovelars lost their labialization and became plain velars (*k'*, *k*, or *g*, according to §§16.13–15 below), as in *elik'* 'he left' < **elikwet*. The palatals became fricatives or affricates – specifically, **k̑* > *s*, **ĝ* > *c*, and **ĝh* > *j*: **dek̑n* 'ten' > *tasn* (with unexplained *-a-*); **ĝenh₁-o* 'race, kind' > *cin*; **bh₂ĝh₂-* 'high' > *barjr* (cp. Eng. *burg*). It is possible that *z* is another outcome of **ĝ(h)*, as in *dēz* 'heap, pile' < **dhoiĝho-* and *mēz* 'urine' < **h₃meiĝho-*, but these may also be loanwords from Iranian.

16.13. The three PIE series of plain voiceless, voiced, and voiced aspirated stops are kept distinct in Armenian, as also in Greek, Italic, Indic, and Germanic; but all three shifted. The shift is similar to Grimm's Law in Germanic (see §15.6) and is often called the **Armenian Consonant Shift**. The plain voiceless stops **t* and **k* (with **k* representing both PIE **k* and **kʷ* as per above) became the voiceless aspirates *t'* and *k'* (e.g. pronominal stem **to-* > Arm. *t'e* 'that'; **elikwet* 'he left' >

elik'). Exceptional was **p*, which only very rarely became the expected *p'* (e.g. *p'etowr* 'feather' < **pet-* 'fly'). Usually it became *h* or completely disappeared at the beginning of a word, as in *howr* 'fire' < **peh₂ur* (cp. Gk. *pŭr* 'fire') and *otn* 'foot' < accusative sing. **pod-n* (cp. Gk. *pod-* and also §16.32 below); and word-internally **p* became *w*, as in the conjunction *ew* 'and' < **epi* (cp. Gk. *epi* 'upon'). For the development of **t* word-internally, see §16.31.

16.14. The voiced stops **b* **d* **g* (from both **g* and **g^w*) became the voiceless stops *p* *t* *k*, as in *tam* 'I give' < **deh₃-* (cp. OCS *damī* 'I will give'), and *kov* 'cow' < **g^wou-* (cp. Eng. *cow*).

16.15. The voiced aspirates **bh* **dh* **gh* (from both **gh* and **g^wh*) become voiced stops. Examples include *berem* 'I carry' < **bber-e-*; *d-nem* 'I place' < **dhh₁-ne-* (from the zero-grade of **dheh₁-* 'place, put', cp. Skt. *dhā-*); and *mēg* 'cloud' < **h₂meigh-* (cp. Ved. *meghā-* 'rain'). The outcome of **gh* (meaning both PIE **gh* and **g^wh*) is complicated by a palatalization rule that changed it to *j* before *e* or *i*, as for example in *jerm* 'heat' from **g^whermo-* (cp. Gk. *thermós*). Word-internally, **bh* was weakened like **p* to *w*, as in the suffix *-awor* 'bearing, having' discussed below in §16.41 that is ultimately from PIE **bhorō-*.

16.16. Liquids and nasals. The liquids and nasals stayed mostly unchanged: 1st sing. present-tense ending *-m* < **-m(i)*; *hing* 'five' < **penk^we-*; *sterj* 'sterile' < **ster-*; *lizem* 'I lick' < **leigh-*. In some positions, especially before other consonants, the *l* was pronounced as a 'dark' or velar *l*, as in *kalmi* 'oak'. Nasals disappeared word-finally after vowels, hence the endingless accusative sing. of words like *get* 'river' < **yed-om* (from the same root as Eng. *wet*). For the initial *g-* in this word, see §16.18.

16.17. The syllabic liquids and nasals developed an *a* before them, as in *angorc* 'inactive' < **ŋ^worǵ-* and *arbi* 'I drank' < **sybh-* (cp. Lat. *sorbere* 'to soak up'). Word-finally, the syllabic nasals became *-n*, as in the numerals *ewt'n* 'seven' and *tasn* 'ten' from **septm* and **dek^m*. Compare also §16.32 below.

16.18. Glides. The PIE glides are not preserved as such. PIE **j* mostly disappeared, as in *erek* 'three' < **trejēs*, but after resonants it became *j*: *jñem* 'I wipe, clean' < **g^when-je-* (root **g^when-* 'bear'), *olj* 'entire' < **ol-jo-*. The other glide, **u*, became *g* word-initially, via an intermediate stage **gu*. This is a common, albeit strange-looking, change; compare the same development in British Celtic, §14.54, and East and North Germanic, §15.41. Examples include *gorc* 'work' < **uorǵo-* (cp. Eng. *work*) and *gitem* 'I know' < **ueid-* (cp. Ved. *véda* 'I know'). Between vowels, however, *u* is the result, as shown most dramatically by the phrase *ger i veroy* 'above and beyond', literally 'above on top', where both *ger* and *ver(-)* continue an earlier **uer-* (formed from **uper* 'above' after the **-p-* was lost).

16.19. Consonant clusters containing **u* show similar developments, some of them quite unusual. The cluster **su* became *k'*, as in *k'oyr* 'sister' from **suesor-*; presumably what happened here is that **su* became **hu* (cp. §16.23 below), which was essentially a voiceless labiovelar glide like the *wh-* in some pronunciations of Eng. *what*; this then developed into a voiceless aspirated velar stop *k'* parallel to the development of voiced **u* into voiced *g* above. Similarly, **u* developed into *k* next to the voiceless stop **k'* in the cluster **ku*, which became *sk*, as in *skownd* 'little dog' < **kuon-t-* (cp. Eng. *hound*).

Especially famous (or infamous) in the annals of IE phonology is the Armenian outcome of PIE **du*: it became *rk*, as in the word for 'two', *erkow* (the *e-* is a later prothetic vowel). While we cannot fully reconstruct all the intermediate stages of this change, it is clear that the velar *k* is the outcome of the glide, as above, and the

r is a rhotacized continuation of the *d*. The change is fully regular, and we have several other examples of it: *erkar* 'long' < **duēh₂ro-* (cp. Doric Gk. *d(w)ārōs* 'long'); *erknč'im* 'I fear' (earlier **erki-nč'im* < **dui-n-sk-*, cp. Gk. perfect (dé)-*d(w)i-men* 'we are afraid'); and *erkn* 'birth-pangs' < **h₁duon-*.

16.20. Laryngeals. Laryngeals, when vocalized, turn into *a* in Armenian, as in *hayr* 'father' < **pater-* < **ph₂ter-*; this is the same development seen in Italic, Celtic, and Germanic. In longer words, sometimes they were vocalized and sometimes not: contrast *dowstr* 'daughter', without a reflex of the laryngeal (< **dhugh₂ter-*, cp. Gk. *thugátēr*), with *arawr* 'plow', which has a reflex (< **h₂erh₂tro-*, cp. Gk. *árottron*).

16.21. Armenian shares with Phrygian and Greek the phenomenon of vocalizing word-initial laryngeals before consonants, usually (but not always) as *a-*, as in *aganim* 'I spend the night' (ultimately from **h₂ues-*, cp. Gk. *á(w)esa* 'I spent the night', Hitt. *huiš-* 'live'), *ayr* 'man' < **aynr-* < **aner-* < **h₂ner-*, and *erek* 'evening' < **h₁reg^wos* (cp. Gk. *érebos* 'darkness of the underworld'). But reflexes of word-initial laryngeals are not preserved as consistently as in Greek.

16.22. Some have claimed that Armenian preserves word-initial laryngeals before vowels as *h*, similar to Anatolian (see §9.5). This claim is based on a number of tantalizing forms such as *haw* 'grandfather' (PIE **h₂eu^h₂os*, cp. Hitt. *huḫḫaš* 'grandfather'), *hanem* 'I draw out' (**h₂en-*, cp. Hitt. *ḫan-* 'scoop'), and *hot* 'odor' (**h₃ed-*). On the other hand, such an *h* is lacking in other words where one might expect it, as *ost* 'branch' (< **h₃ezd-*, cp. Hitt. *ḫašduir*) and *oror* 'gull' (if from **h₃er(o)n-*, cp. Hitt. *ḫaran-* 'eagle'). A rather grave difficulty is the fact that *h*'s were not infrequently added to vowel-initial words, leading to doublets like *ogi* 'breath' ~ *hogi* 'spirit' and *arbil* 'to drink' ~ *harbil* 'to be intoxicated'. The interpretation of forms like *haw* is therefore disputed.

16.23. Sibilant **s*. PIE **s* usually disappeared word-initially and between vowels, as in the words *at* 'salt' (< **sal-*) and *k'oyr* 'sister' (< **h₂ue^r₂r* < **suesor-*). But sometimes it became *h* initially, as in *bin* 'old' < **seno-*. At the end of words, it appears that the outcome is (oddly) often *k'*, but as this outcome is only found in plurals, we do not know if the *-k'* comes directly from **-s* or is due to some morphological analogy. Nevertheless, the correspondences are rather striking: compare the nomin. pl. *otk'* 'feet' with Gk. *pódes*; the 1st pl. verbal ending *-mk'* with Lat. *-mus*, Ved. *-mas*; and note also the word for 'three', *erek* < **trejes*. In at least one environment, final **-s* remained, namely the accusative plural from **-ns*, as in *gets* 'rivers' < earlier **uedons*.

VOWELS

16.24. The short vowels usually stayed intact: *ac-em* 'I lead', cp. Lat. *ag-ō* 'I lead'; *cer* 'old man', cp. Gk. *gérōn* 'old man'; *elik'* 'he left', cp. Gk. *élipe* 'he left'; *hot* 'smell', cp. Lat. *odor*; *dowstr* 'daughter' (*ow* = *u*), cp. Gk. *thugátēr* 'daughter'. In unaccented syllables *i* and *u* were deleted (see §16.28 below).

16.25. The long vowels became short. The long mid vowels **e* and **o* were raised to *i* and *ow* (= *u*): *mi* (negative used in prohibitions), cp. Gk. *mé*; *towrk'* 'gift', cp. Gk. *dōron*. Long **ā* was shortened to *a*, as in *awr* 'day', cp. Gk. (Doric) *āmar*. Long **ī* was shortened to *i*, as in *k'san* 'twenty' < earlier **k'isan* < **uik^mmti*, and **ū* was probably shortened as well, as in *jowkn* 'fish' if from **dhghū-*.

16.26. Diphthongs. The PIE diphthongs **ai* and **au* remain unchanged, as in *ayc* 'goat' (cp. Gk. *aig-*) and *awl* 'place to spend the night' (cp. Gk. *aūlis*). The diphthongs **eu* and **ou* both became *oy*: *loys* 'light' (**leuk-* or **louk-*). Finally, **ei* and

*oi became *ē, as in the 3rd sing. aorist *elēz* 'he licked' (< **eleighet*) and *tēz* 'heap, pile' (< **dhoigh-o-*, cp. Gk. *toikhos* 'wall').

STRESS AND LOSS OF FINAL SYLLABLES

16.27. At some point in the prehistory of Armenian, the mobile PIE accent was fixed on the penultimate syllable. After this, most vowels in final syllables were lost, and – usually but not always – final consonants. Thus for example the PIE imperfect **ebheret* 'he was carrying' became Armenian (aorist) *eber*. The loss of final syllables resulted in the stress now being word-final, since the old penultimate syllables became the new final syllables. (The same thing happened in the history of French to give the word-final accent that Modern French has; a word like *interdît* 'forbidden' comes from Latin *interdictus*.)

16.28. The stress on final syllables resulted in weakening and syncope of certain non-final syllables, and led to a series of vowel alternations depending on the position of the stress. Underlying *i* and *ow* (*u*) disappear outright when not stressed or word-initial: thus compare *gir* 'letter' with *grem* 'I write' (< **girem*) and *k'own* 'sleep' with (genitive) *k'noy* 'of sleep' (< **k'ownoy*). Other alternations are found as well: *mēg* 'cloud' (earlier **meig*) ~ *migamac* 'foggy'; *loys* 'light' ~ *lowsawor* 'luminous'; *matean* 'book' ~ *matenagir* 'writer'. (*Migamac* and *lowsawor* show that when *i* and *ow* result from reduction of a diphthong, they remain even when unstressed.)

Morphology

16.29. Armenian has undergone moderate simplification of PIE inflectional morphology. A total of seven cases – most of the PIE inventory – are distinguished in nouns and pronouns (nominative, accusative, genitive, dative, instrumental, locative, and ablative), although no single noun or pronoun distinguishes all seven. The dual has been lost everywhere. The verbal system of Classical Armenian is quite similar to that of PIE, but interestingly few of the actual forms are inherited. The present and aorist have survived, while the perfect has disappeared, as has the imperfect (although several individual imperfects were transferred over to the aorist category and survive as aorists, such as *eber* above). A new imperfect was created, as were a present and aorist subjunctive. Mediopassive inflection is distinguished from active inflection by a stem change, and by partially different endings in the aorist. Only one participle is found.

NOUNS AND PRONOUNS

16.30. Nouns were declined in vocalic and consonantal stems. Four vocalic stems, in *o*, *a*, *i*, and *ow* (*u*), are found, exemplified below by the declensions of *get* 'river', *am* 'year', *ban* 'word', and *zgest* 'clothing':

	sg.	pl.	sg.	pl.	sg.	pl.	sg.	pl.
N	<i>get</i>	<i>getk'</i>	<i>am</i>	<i>amk'</i>	<i>ban</i>	<i>bank'</i>	<i>zgest</i>	<i>zgestk'</i>
A	<i>get</i>	<i>gets</i>	<i>am</i>	<i>ams</i>	<i>ban</i>	<i>bans</i>	<i>zgest</i>	<i>zgests</i>
GD	<i>getoy</i>	<i>getoc'</i>	<i>ami</i>	<i>amac'</i>	<i>bani</i>	<i>banić'</i>	<i>zgestow</i>	<i>zgestowc'</i>
L	<i>get</i>	<i>gets</i>	<i>ami</i>	<i>ams</i>	<i>bani</i>	<i>bans</i>	<i>zgestow</i>	<i>zgests</i>
Ab	<i>getoy</i>	<i>getoc'</i>	<i>amē</i>	<i>amac'</i>	<i>banē</i>	<i>banić'</i>	<i>zgeste</i>	<i>zgestowc'</i>
I	<i>getov</i>	<i>getovk'</i>	<i>amaw</i>	<i>amawk'</i>	<i>baniw</i>	<i>baniwk'</i>	<i>zgestow</i>	<i>zgestowk'</i>

The genitive and dative are not distinguished in nouns, only in pronouns. The plural *-k'* and the accusative plural *-s* have already been discussed. The instrumental contains a *-v-* or *-w-* that continues the PIE instrumental ending **-bhi-*; this ending survives more clearly in such old athematic forms as *jerb-a-kal* 'taken with the hands = prisoner' < **ġhesy-bhi-*. The genitive and ablative plural ending *-c'* is of uncertain origin.

16.31. The declension of consonant stems is similar, although some show unexpected alternations in the stem. For example, *hayr* 'father' has a genitive *hawr*. The difference in vocalism is due to different treatments of word-internal **t:* before a front vowel this stop became the glide *y* (**hayir* < **hatir* < **ph₂tēr*), while before back vowels and in the consonant cluster **-tr-* it became the other glide *w*, hence *hawr* < **hatros* < **ph₂tros*. Another kind of stem-alternation in a consonant stem can be exemplified by *p'ok'r* 'small', pl. *p'ok'ownk'*; it is ultimately a *u*-stem, into which an *-r* has intruded in the nominative and accusative singular and an *-n-* in the plural. This is reminiscent of the PIE *r/n*-stems, although it is not fully clear how this Armenian paradigm came to be. A true *r/n*-stem declension does not exist in Armenian, but a trace remains in the *r/n*-alternation seen in *howr* 'fire' vs. *hn-oc'* 'oven' (cp. §6.31).

16.32. A small group of nouns, principally *otn* 'foot' and *jein* 'hand', end in an *-n* that actually continues the old accusative singular ending **-nī*. The loss of this ending in the productive accusative singular of consonant stems is therefore not due to sound change, but to morphological analogy with the vocalic stems, where all endings in **-Vm* were lost.

16.33. Pronouns. Of interest is the system of deictics ('pointing' words such as demonstrative pronouns, adjectives, and adverbs). Unlike English, which has only a two-fold distinction between *this* and *that* and between *here* and *there*, Armenian, like Latin and certain other ancient IE languages, has a three-fold distinction corresponding to the three persons. Unlike Latin, though, the three-way distinction is systematically carried out throughout the system of demonstrative pronouns and adverbs. Armenian thus has three definite articles, the suffixes *-s*, *-d*, and *-n* (ultimately < PIE **ki-*, **to-*, and **eno-lono-*), meaning roughly 'this (by me)', 'that (by you)', 'that (over by him/her/them)'. From these are formed three demonstrative pronouns *ays* 'this', *ayd* 'that', and *ayn* 'that (over there)'; three corresponding anaphoric pronouns *as*, *da*, and *na* basically function as third person pronouns but again depending on how near to the speaker the person or thing referred to stands (either physically or metaphorically). Armenian also has not one but three pronouns meaning 'the same', *soyn*, *doyn*, *noyn*; three locational adverbs for 'here' or 'there', *ast*, *aydr*, and; and three interjections meaning 'behold', *awasik*, *awadik*, *awanik*.

The definite article suffixes can be attached not only to nouns, but to any part of speech, so long as it is part of the phrase or clause that is being marked definite. (Compare from English the ability of the possessive suffix *-s* to be added to any part of speech so long as it comes at the end of the noun phrase being marked for possession, as in *the woman I saw yesterday's coat*.) Thus in *zor oc'd kayel ē k'ez xawxel* 'which it is not seemly for you to relate' (Agat'angelos 68), the article *-d* is attached to the negative *oc'* 'not'.

VERBS

16.34. There were two verb stems, a present stem and an aorist stem, and they expressed not only an opposition in tense (non-past vs. past) but also an opposition

in aspect (imperfective vs. perfective; see §5.10). While these basic categories are inherited, the personal endings have undergone significant change; and a third stem, the perfect, has been entirely lost. Some historical details of the personal endings will be taken up in the sections to follow.

The present stem is used to form the present and imperfect indicative, the present subjunctive, the prohibitive, and the infinitive. The aorist stem is used to form the aorist indicative and subjunctive, the imperative, and a passive participle.

16.35. Present classes. There are four classes of presents, with stem vowels *a*, *e*, *i*, and *ow* (*u*). The following verbs will illustrate the present and imperfect: *orsam* 'I hunt', *orsayi* 'I was hunting'; *grem* 'I write', *grei* 'I was writing'; *hayim* 'I look', *hayei* 'I was looking'; and *argelownm* 'I hinder', *argelowi* 'I was hindering'.

	Present				Imperfect			
sg. 1	<i>orsam</i>	<i>grem</i>	<i>hayim</i>	<i>argelownm</i>	<i>orsayi</i>	<i>grei</i>	<i>hayei</i>	<i>argelowi</i>
2	<i>orsas</i>	<i>gres</i>	<i>hayis</i>	<i>argelows</i>	<i>orsayir</i>	<i>greir</i>	<i>hayeir</i>	<i>argelowir</i>
3	<i>orsay</i>	<i>grē</i>	<i>hayi</i>	<i>argelow</i>	<i>orsayr</i>	<i>grēr</i>	<i>hayēr</i>	<i>argeloyr</i>
pl. 1	<i>orsamk'</i>	<i>gremk'</i>	<i>hayimk'</i>	<i>argelownmk'</i>	<i>orsayak'</i>	<i>greak'</i>	<i>hayeak'</i>	<i>argelowak'</i>
2	<i>orsayk'</i>	<i>grēk'</i>	<i>hayik'</i>	<i>argelowk'</i>	<i>orsayik'</i>	<i>greik'</i>	<i>hayeik'</i>	<i>argelowik'</i>
3	<i>orsan</i>	<i>gren</i>	<i>hayin</i>	<i>argelown</i>	<i>orsayin</i>	<i>grein</i>	<i>hayein</i>	<i>argelownin</i>

The PIE primary personal endings (sing.) **-mi* **-si* **-ti*, (pl.) **-mes(-)* **-tes(-)* **-nti* can be seen shining through the Armenian present-tense endings. The preservation of the *-s* in the 2nd singular is probably due to influence from the 2nd singular of the verb 'to be', which is *es* < **ess* < **(h₁)es-si* (remade from **h₁esi*; §3.37). The 3rd singular *-y* is one of the regular developments of intervocalic **t* (§16.31): thus *-ay* comes from **-a-ti*, *-ē* (< **-ei*) from **-e-ti*, etc. The imperfect endings are modeled on the imperfect of the verb 'be', whose paradigm is *ei eir ē* in the singular, and *eak' eik' ein* in the plural. These forms in turn have a complicated history, but the *-i-* that characterizes many of them is probably a continuation of **ēs-* (**e-h₁es-*), the old imperfect stem of **h₁es-* 'be'. The 3rd sing. imperfect ending *-r* is an old middle ending; apparently Armenian generalized the middle ending to the 3rd singular of this tense.

16.36. Origin of the present classes. These four classes of present stems continue most of the PIE present formations; the general situation is not unlike that of Italic (see §13.13). Presents in *-em* are kind of a scrap-heap: this class includes old thematic presents like *berem* 'I carry' (**bher-e-*) and *acem* 'I drive' (**h₂eg-e-*); **-iē/o-*presents like *injem* 'I wipe, I clean' (< **g^hhen-iē-*); and causative-iteratives like *owtem* 'I eat' (< **h₁ōd-ēiē-*) and *glem* 'I roll' (< **gowlem* < **uōl-ēiē-*). Very productive in Armenian is the nasal suffix *-an-*, which was often used to remake already-existing verbs: *anicanem* 'I curse' is a remodeling of **anicem*, a **-iē/o-*verb (**h₃neid-iē-*, cp. Gk. *ónēidos* 'blame'); *harc'anem* 'I ask' is a remodeling of a **-skē/o-*verb (**pr(k)ske-*, cp. Lat. *poscō* 'I ask'; see §5.31); and *lk'anem* 'I leave' is a remodeling of a nasal-infix verb (cp. Lat. *linquō* 'I leave'). The suffix *-an-* is related to the Greek suffix *-anein*, which was added to many verbs that already had a nasal infix, such as *limp-ánein* 'to leave'.

Other PIE present formations are scattered among the remaining classes, such as root athematic presents (e.g. *bam* 'I speak' < **bheh₂-mi*) and *nu-*presents (e.g. *z-genownm* 'I get dressed' < **ues-nu-*, exactly like Gk. *hénnumi* 'I dress [someone]'; §5.23). The *i-*stem verbs typically continue middles formed with the suffix **-iē-*, as in *meranim* 'I die' (cp. Ved. *mri-yá-te* 'dies', Lat. *mor-ior* 'I die').

16.37. The present passive. Verbs with stem vowel *-e-*, like *berem*, could form a passive in *-im* (so *berim* 'I am carried'), with endings exactly the same as those of the *i*-stem verbs like *hayim*. This passive conjugation is derived from the PIE suffix **-je-* (§5.29), which also furnishes the passive conjugation in Indo-Iranian.

16.38. The aorist. The Armenian aorist for the most part is a continuation of the PIE imperfect. There are two types of aorist stems: an unextended root (e.g. 3rd singulars *e-ber* 'carried', from an old IE imperfect **ebheret*, and *elik* 'left', from an old IE thematic aorist **elik^wet*, equivalent to Gk. *ēlipe*); and a root extended with a suffix *-c'* (e.g. 3rd sing. *gorcec'i* 'he made' from the present *gorcem* 'I make'), usually derived from **-skē-*, which was used in Greek and Anatolian to form iterative past tenses (see §5.31). Unlike the present, the aorist has separate active and mediopassive personal endings. Some examples of the conjugations are provided below by the active aorists *hani* 'I drew out' and *orsac'i* 'I hunted', and the mediopassive aorists *nstay* 'I sat' and *hayec'ay* 'I looked':

	Aorist indicative active		Aorist indicative mediopassive	
sg. 1	<i>hani</i>	<i>orsac'i</i>	<i>nstay</i>	<i>hayec'ay</i>
2	<i>haner</i>	<i>orsac'er</i>	<i>nstar</i>	<i>hayec'ar</i>
3	<i>ehan</i>	<i>orsac'</i>	<i>nstaw</i>	<i>hayec'aw</i>
pl. 1	<i>hanak'</i>	<i>orsac'ak'</i>	<i>nstak'</i>	<i>hayec'ak'</i>
2	<i>hanëk'</i>	<i>orsac'ëk'</i>	<i>nstayk'</i>	<i>hayec'ayk'</i>
3	<i>hanin</i>	<i>orsac'in</i>	<i>nstan</i>	<i>hayec'an</i>

As can be seen from the form *ehan* (and *eber* cited earlier), a prefix called the augment (*e-*), identical to the augment found in Greek, Indo-Iranian, and Phrygian (§5.40), appears when the form would otherwise be a monosyllable. The active aorist endings are in part a continuation of the PIE imperfect; the 3rd singular, for example, continues the imperfect thematic ending **-et*. The *-w* at the end of the 3rd singular aorist passive is a faithful descendant of the PIE past-tense middle ending **-to* (for the change of **-t-* to *-w-* see §16.31). Similarly, the 3rd pl. *-an* is from **-nto*; this became **-anto* and then *-an* after the loss of final syllables, and the *-a-* in this form is thought to have spread throughout the paradigm and to be the source of the *-a-* in the other persons and numbers.

16.39. The subjunctive. The subjunctive functioned both as a future tense and as an optative and conditional. It was formed with a suffix *-ic'* which is usually taken to be from PIE **-iske-*, although this is not universally accepted. Its endings are complex, and not explained in all their particulars. One form that is generally agreed upon is the 1st singular, as in *argel-ic'* 'I will hinder', where the ending *-ic'* is taken to be from **-iskō* (**-iskoh₂*), with the original ending **-ō* (< **-oh₂*) not replaced by **-m(i)* as in the present indicative.

16.40. The participle. Armenian verbs have only one participle, in *-eal*, formed typically from the aorist stem and having passive meaning, such as *greal* 'written'. This goes back ultimately to the somewhat rare PIE participial ending **-lo-* (§5.56), also found in Slavic, as in OCS *bi-lŭ* 'beaten'.

COMPOUNDING

16.41. Armenian is very fond of using compounding to form new words; interestingly, it almost never uses prefixes, a feature shared with PIE, though not necessarily

inherited from it: to judge by the presence of a few archaic prefixed forms, prefixation may have once been more common. Two examples are *anker* 'companion' (< **and-ker-*, literally 'with-eat(er), messmate') and *tkar* 'weak' (< **ti-kar* 'without power'; **ti* is cognate with Lat. *dē* 'from, without', as in *dē-bilis* 'without strength, incapable'). Much preferred was suffixation, and Armenian possessed a wealth of suffixes. Some examples of suffixed forms are *lows-awor* 'luminous' (*loys* 'light' + *-awor* 'bearing, having' < PIE **bhoro-*), *xalat-arar* 'peaceful' (*xalat* 'peace' + *-arar* 'doing', from *ainem* 'I do'), and *k'ahanay-owt'iwn* 'priesthood' (*k'ahanay* 'priest' + *-owt'iwn* '-hood, -ship', related to the Latin abstract nouns in *-tiōn-*). Perhaps most familiar to English speakers is the patronymic suffix *-ean* (ultimately from PIE **-ijo-*, see §6.74) meaning 'son of', as in *Aram-ean* 'son of Aram' or *Simownean Yowda* 'Judas son of Simon' in the Armenian Bible (John 6:71). This appears transcribed from Modern Armenian into English as *-ian* in such names as *Khatchaturian*, *Hagopian*, etc.

Armenian text sample A

16.42. The Birth of Vahagn, one of the native pre-Christian poems quoted in Movsēs Xorenac'i 1.31. Vahagn (or Vahevan), from Parthian *Varθagan*, is in origin the Iranian hero *Vərəθraγna-*, equivalent etymologically to Ved. *Vṛtra-ghn-* 'slayer of Vṛtra' (the serpent and monstrous adversary of the chief god Indra). This legend is not Iranian in origin, but probably Anatolian.

- | | | |
|---|---|---|
| 1 | <p>erknēr erkin erknēr erkir
 erknēr ew covn cirani
 erkn i covown ownēr
 zkarmrikn elegendik.</p> | <p>Heaven was in labor, Earth was in labor,
 the purple sea was also in labor.
 In the sea labor pains held
 the little crimson reed.</p> |
| 2 | <p>and elegend p'ol cowx elanēr
 and elegend p'ol boc' elanēr
 ew i boc'oyñ vazēr
 xarteaš patanekik.</p> | <p>Along the reed's stalk rose smoke;
 along the reed's stalk rose flame;
 and from the flame leapt up
 a golden-haired little boy.</p> |
| 3 | <p>na howr her ownēr
 boc' ownēr mawrows
 ew ač'kownk'n ein aregakownk'.</p> | <p>He had hair (of) fire;
 he had a beard (of) flame;
 and his eyes were little suns.</p> |

16.42a. Notes. 1. *erknēr*: 'was in labor', from *erkn* 'labor pains' (in line 3), stem *erkown-*, PIE **h₁d-ye-* (for **dū > rk*, see §16.19), also in Gk. *odūnē* 'pain', a derivative of **h₂ed-* 'eat' in its presumed original meaning 'bite'. A parallel suffixed form in *-*ye-* is represented by the Anatolian words for 'evil', Luv. *attinwal-* and Hitt. *idālu-*. *erkin*: 'heaven', etymology unclear. *erkir*: 'earth', etymology likewise unclear, but forming an obvious pair with *erkin*. *ew*: 'and, also', < **epi* 'upon'; §16.13. *covn*: 'the sea'; -*n* is the cliticized definite article (seen also below in *covown*, *zkarmrikn*, *boc'oyñ*, etc.). It has been suggested that this is one of the Urartian borrowings into Armenian (§16.2), but it may also be cognate with Lydian *koṣul* 'water'. The locative is *covow*, seen below, showing the word to be a *u*-stem. *cirani*: 'purple'. *i*: 'in' (plus locative); when followed by the ablative it means 'from', as in the second stanza. *ownēr*: 'had, held', 3rd sing., usually taken to be from a nasal present **ōp-ne-* from the same root as in Ved. *āpnōti* 'gets, takes'. *zkarmrikn*: 'the crimson', from *karmir* 'red' (cp. Hebrew *karmil*, perhaps from Persian and ultimately related to Eng. *crimson*). The prefix *z-* marks the word

as direct object; the prefix was originally a preposition but developed somewhat like Spanish *a* (from Lat. *ad* 'to') to mark direct definite objects. *elegnik*: 'little reed', diminutive of *elegn* 'reed'. Note that in a definite noun phrase, either the adjective or the noun can host the definite article; in this phrase, it is the adjective.

2. *ənd*: 'along', from **h₂enti* 'facing, against'. *elegan*: 'reed', *n*-stem genit. sing. *p'ot*: 'stalk'. *cowx*: 'smoke'. *elanēr*: 'ascended, rose'. *boc'*: 'flame'. The ablative is *boc'oy* in the next line, with a suffixed definite article *-n*. *vazēr*: 'leapt'. *xarteaš*: 'golden-haired'. *patanekik*: 'little youth', containing the same diminutive suffix *-ik* as *elegnik* above.

3. *na*: 'that one, he'. *howr*: 'fire', PIE **peh₂ur* (see §§16.13 and 31). *her*: 'hair'; not related to the English word. *mawrows*: 'beard', accus. pl. (but singular in meaning), from PIE **smōkrū-*, cognate with Ved. *śmāśru-*, Lith. *smākras* 'chin', and other forms. The Armenian word illustrates the loss of **s-* before nasals, a regular Armenian change. The cluster *-kr-* developed to *-ur-* just as **-tr-* did (§16.31). *ač'kownk'n*: 'the little eyes', from earlier **ač'-ikounk'-n*, the definite (*-n*) plural (*-ownk'*, with the same intrusion of *-own-* into the plural paradigm as seen in the adjective *p'ok'r* in §16.31) of a diminutive (with suffix *-ik-*) of *akn* 'eye', PIE **h₂ek'*. (cp. Lat. *oculus*). *cin*: 'were', 3rd pl. imperfect. *aregakownk'*: 'little suns', from *aregakn*, literally 'sun-jewel', from *areg*, *arev* 'sun' (cognate with Skt. *ravis* 'sun') plus *akn* 'jewel', perhaps the same word as *akn* 'eye'.

Armenian text sample B

16.43. From the Armenian Bible; Mark 14:17–21. The English translation is adapted from the *New International Version*.

17 Ew ibrew erekoy elew gay erkotasaniwk'n handerj: 18 Ew ibrew bazmec'an ew deŕ owtein, asē Yisows. amēn asem jez, zi mi omn i jēnĭ matneloc' ē zis, or owte isk ənd is: 19 Ew nok'a sksan trtmel ew asel mi əst misjē. mit'e es ic'em, ew miwsn mit'e es ic'em: 20. Na patasxani et ew asē c'nosa mi yerkastasanic' ayti, or mxeac' ənd is skawarəkd: 21. Ayl ordi mardoy ert'ay orpēs ew greal ē vasn nora. bayc' vay ic'ē mardoyn aynmik yoyr jeŕs ordi mardoy matnesc'i. law ēr nma t'ē c'ēr isk cneal mardn ayn.

17 When evening came, Jesus arrived with the Twelve. 18 While they were reclining at the table eating, he said, "I tell you the truth, one of you will betray me – one who is eating with me." 19 They were saddened, and one by one they said to him, "Surely not I?" 20 "It is one of the Twelve," he replied, "one who dips bread into the bowl with me. 21 The Son of Man will go just as it is written about him. But woe to that man who betrays the Son of Man! It would be better for him if he had not been born."

16.43a. Notes (selective). 17–18. *ew*: 'and', etymologically 'thereupon' or the like, from PIE **epi* 'upon'. *erekoy*: 'evening', from **h₂reg^{wo}*- (cp. Gk. *érebos* 'darkness'). *gay*: 'comes', 3rd sing. of *gam* 'I come', PIE **ǵādh-* (also the root of Lat. *uadere* 'to come, go', as in *Quō uādis?* 'Whither goest thou?'). *erkotasaniwk'n*: 'group of twelve', instr. sing. as object of *handerj* 'together with', from *erko* 'two' and *tasn* 'ten'. *owtein*: 'were eating', 3rd pl. imperfect of *owtem* 'I eat', from PIE **h₂ōd-éje-*, a lengthened-grade causative-iterative from a Narten root (see §5.20). *asē*: 'says', PIE **h₂eĝ-* 'say', the root of Gk. *ē* 'he said' and Lat. *ait* 'he said'. *mi*: 'one', PIE **smi-*, cp. Gk. (fem.) *mía* < **smi(l)h₂*. *ənd*: 'with', see the Notes to line 2 of the preceding text.

19–21. es: 'I', PIE **eg-*, ic'em: 'I will be', present subjunctive of *em* 'I am'. *patasxani* et: '(he) answered'; phrase made of the noun *patasxani* 'answer' and the aorist of the verb *tam* 'I give', PIE **deh₃-*. c'nosa: 'to them', preposition c'- 'to' plus accus. pl. of *na*, 3rd person pronoun. mardoy: 'of man', genitive of *mard*, either an Iranian borrowing (cp. Modern Persian *mard* 'man') or directly from PIE **mrto-* 'mortal' (also the source of the Iranian). greal ē: 'it is written', periphrastic passive with the participle of *grem* 'I write' plus *ē*, 3rd sing. present of *em* 'I am'. *matnesc'i*: 'will betray', 3rd sing. aorist subjunctive of *matnem* 'I betray', literally 'I finger, point a finger at', from *matn* 'finger'. *cneal*: 'born', participle of *cnem* 'I am born', PIE **ǵenb₁-*.

For Further Reading

Not much of any value has been written on the history of Armenian in English. The classic comparative grammar is still Meillet 1936, amazingly thorough in spite of its brevity. The grammar is partly a distillation of decades of pathbreaking scholarly work that Meillet devoted to Armenian, which can be found in the masterly articles collected in Meillet 1962–77. Newer comparative grammars are Schmitt 1981, useful for the reader with some IE background, and Lamberterie 1992. A reference grammar of Classical Armenian in English is Godel 1975. An excellent, though highly technical and not uncontroversial, study of the history of the Armenian verb is Klingenschmitt 1982. The only etymological dictionary not in Armenian dates from the nineteenth century and is badly in need of revision: Hübschmann 1895–7. The question of the relationship between Greek and Armenian has recently received a detailed investigation (if not resolution) in Clackson 1994. The Armenian noun has recently received detailed morpho-historical treatment in Olsen 1999.

For Review

Know the meaning or significance of the following:

Heinrich Hübschmann
Mesrob

Armenian Consonant Shift
deictic

Exercises

- What are the Armenian outcomes of the following PIE sounds or sequences of sounds? Some may have more than one answer.

a	*bh	c	*t	e	*k	g	*y	i	*dy	k	*kʷ	m	*s	o	*ʃ
b	*j	d	*p	f	*j	h	*d	j	*ǵh	l	*ð	n	*sy	p	*eu

- Briefly explain the history or significance of the following Armenian forms:

a	<i>kov</i>	f	<i>otn</i>	k	<i>gorc</i>
b	<i>k'oyr</i>	g	<i>haw</i>	l	<i>tasn</i>
c	<i>ger i veroy</i>	h	<i>erkow</i>	m	<i>ayr</i>
d	<i>-mk'</i>	i	<i>jerbakal</i>	n	<i>lowsawor</i>
e	<i>hnoc'</i>	j	<i>bam</i>	o	<i>acem</i>

3 Consider the following PIE roots and their Armenian descendants:

* <i>ǵenh₁-</i>	<i>cin</i>	'birth'
* <i>penk^we</i>	<i>hing</i>	'five'
* <i>sen-</i>	<i>hin</i>	'old'
* <i>ǵer-</i>	<i>cer</i>	'old'
* <i>syeks</i>	<i>vec</i>	'six'

- Formulate a rule that accounts for the changes that affected PIE **e* in the data provided.
 - Given that the verb 'to engender' derived from *cin* 'birth' is *cnel*, which happened first, your rule in (a) or the sound change in §16.28?
- Among the many Iranian loanwords in Armenian are *partēz* 'paradise' (cp. Avestan *pairi.daēza-* 'enclosure') and *dast* 'hand' (cp. Middle Persian *dast*). Based on your knowledge of Armenian consonantal sound changes, which of these is the earlier borrowing into Armenian? Explain your answer.
 - How do forms like *barjr* and *dēz* (§16.12) provide evidence that Grassmann's Law was a post-PIE sound change? What other branch(es) that you have seen so far provide(s) comparable evidence? Explain your answer.
 - The preservation of the short vowels intact into the historical period (§16.24) is shared by what other branch(es) of IE that you have seen so far?
 - How does the Armenian class of verbs with stem-vowel *e* differ historically from the Italic 3rd conjugation (also with stem-vowel *e*)?
 - What was the fate of the interdental fricative [θ] (written *ð*) in Parthian loanwords into Armenian, as exemplified below?

aparan-k' 'house, palace' < Parth. *apaðan*
awrēn 'law, right' < Parth. *aβðēn*
varagoyr 'curtain' < Parth. *barayðð*

PIE Vocabulary VIII: Material Culture and Technology

- **teks-* 'fashion, construct': Hitt. *takkešzi* 'puts together', Ved. *tāṣṭi* 'fashions', Gk. *téktōn* 'carpenter', Lat. *texō* 'I weave'
- **k^wek^w/lo-* 'WHEEL': Hitt. *kugullaš* 'donut', Ved. *cakráṃ*, Gk. *kúklos*
- **dom-* 'house': Ved. *dámas*, Gk. *dómos*, Lat. *domus*, OCS *domŭ*
- **dh₁uer-* 'door': Av. *duuar-* 'gate', Gk. *thur-*, Lat. *forēs*, OCS *dvŕi*
- **uebh-* 'WEAVE': Ved. *ubhnāti* 'ties together', Gk. *huphaínō* 'I weave', Toch. B *wāp-*
- **s(i)uH-* 'SEW': Ved. *syūtá-* 'sewn', Lat. *suō* 'I sew', Lith. *siúti* 'to sew'
- **ues-* 'clothe, WEAR clothes': Hitt. *waššezi* '(he) clothes', Ved. *váste* 'gets dressed', Gk. *héstai* 'gets dressed', Lat. *uestis* 'clothing'

17 Tocharian

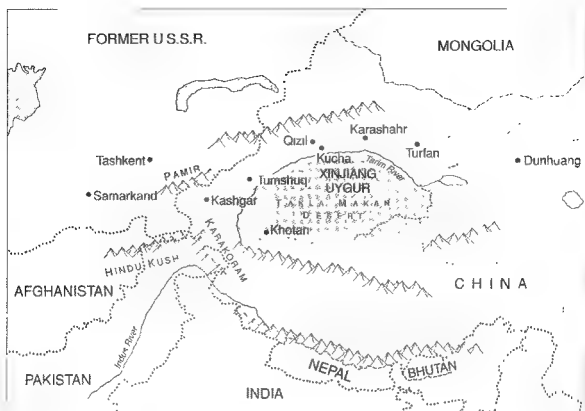
Introduction

17.1. Tocharian, like Hittite, did not come to light until the twentieth century. For about twenty years up until the First World War, the French, Germans, and British undertook numerous archaeological expeditions to Chinese Turkestan (now the Xinjiang Uygur Autonomous Region in western China) and unearthed documents in four previously unknown languages. Two of these turned out to be Middle Iranian (Khotanese and Tumshuqese, see §§11.35 and 44), while the others were the two Tocharian languages. They were recognized as Indo-European already in 1907; their decipherment was greatly aided by the fact that most of the texts were translations (some bilingual) of familiar Buddhist works that had been widely disseminated in Central Asia. The Tocharian documents all date to between the sixth and eighth centuries AD.

17.2. From the lowland region known as the Tarim Basin in eastern Turkestan, in the area around the oases called Qārāšahr (Karashahr) and Turfan, comes Tocharian A, also called East Tocharian, Turfanian, or Agnean (after Agni, the Sanskrit name of Qārāšahr). From that area and farther southwest around Kučā (Kucha) comes Tocharian B, also called West Tocharian or Kuchean. Since Tocharian A is only known from areas where documents in Tocharian B have also been found, it has been suggested that Tocharian A was in fact an already extinct liturgical or poetic language that was kept alive by tradition, and that Tocharian B was the living administrative language. (Of the minority of texts that are not Buddhist translations – including monastic letters, caravan passes, business letters, and graffiti – almost all are in Tocharian B.) One set of Tocharian B texts, called the MQ texts (after the caves of Ming-öi Qizil west of Kučā, where they were found), are written in a variant dialect and may preserve some archaic distinctions in the vowel system; these have still not been adequately studied. A passage from an MQ text is given as one of the text selections at the end of this chapter.

17.3. Who the Tocharians were is still enshrouded in mystery and considerable debate. They left behind no texts about themselves, and it is unclear which of the ethnonyms preserved in contemporary Classical and Central Asian sources refers to them. The designation *Tocharian* is based on the theory, now no longer accepted, that the Tocharian languages were spoken by a Central Asian people called the *Tókharoi* in Greek sources.

We also do not know how and when they, as Indo-Europeans from far to the west, got to such a distant corner of the world: of all the ancient IE languages, Tocharian



Map 17.1 The Tarim Basin

was spoken farthest to the east. Which of the other branches of IE Tocharian is most closely related to is also an unsolved problem; a set of rather archaic features of its verbal morphology have led some researchers to claim that it, like Anatolian, split off from the rest of the family rather early. An earlier view, now abandoned, connected Tocharian with the supposed Italo-Celtic subgroup of IE (§13.5).

17.4. The vocabulary of our Tocharian texts, since they are mostly Buddhist, has been influenced by Sanskrit and Iranian, from which many religious terms were borrowed. (One of them, interestingly, may have made its way into English as the word *shaman*, if ultimately from Tocharian B *šamāne* 'monk', from Sanskrit *śramaṇas*.) Some structural features, such as the large number of cases in the noun (§17.23 below) and the limited stop inventory (only voiceless stops), are not typical of IE languages but are found in Uralic, Turkic, and Mongolian languages of western and central Asia. (Uralic is the language family to which Finnish and Hungarian belong.) It has been suggested that the Tocharians picked these features up from contact with those languages after they migrated eastward.

17.5. Adding to the various mysteries surrounding the Tocharians is the existence of extremely well-preserved mummies in the Takla Makan desert that have striking Europoid features and often red hair; some are nearly 4,000 years old. The mummies were found with tapestries woven in plaids that are similar in weaving style and pattern to tartans from the Hallstatt culture of central Europe, which was ancestral to the Celts (see §14.2); the wool used in weaving the tapestries comes from sheep of European ancestry. Certain cave paintings in the region from later times depict warriors with red hair and other non-Chinese or non-Mongolian features. Whether any of the mummies or the paintings are to be connected with the Tocharians is not known.

Most of the Tocharian that survives is found on fragments of manuscript leaves that were left in Buddhist temples and shrines in the desert as votive offerings. Once left in the shrines, the leaves were picked up by the wind; some landed in the desert sands, where they were buried and eventually dug up by archaeologists. Because of this manner of preservation, we usually have only one leaf from any given Tocharian manuscript.

17.6. With the exception of a few fragments of Tocharian B that are written in Manichean script, Tocharian was written in a modified version of the north Indian Brāhmī alphabet (see §10.57). The same version was used to write the Middle Iranian language Tumshuqese (§11.44). Among the modifications made to the script is the so-called *Fremdvokal* or 'foreign vowel', transliterated as *ä* (discussed in §17.14 below).

From PIE to Tocharian

In the ensuing discussion, forms in Tocharian A and Tocharian B will be labeled simply A and B, and AB will be used to denote a form identical in both languages.

Phonology

CONSONANTS

17.7. Stops. As mentioned above, Tocharian has an unusual stop consonant inventory by IE standards: there were no voiced stops – only *p*, *t*, and *k*. All three manners of stop articulation in PIE – voiceless, voiced, and voiced aspirated – therefore merged into voiceless. This is most clearly seen with the labials, which all became *p*: PIE **prek-* 'ask' > A *prak-*, B *prek-* 'ask' (cp. Lat. *precor* 'I ask, entreat'); PIE **dhub-ro-* 'deep' > A *tpär*, B *tapre* 'high' (cp. Eng. *deep* from full-grade **dheub-*; for the semantics, cp. Lat. *altus* 'high, deep'); PIE **bhrātēr* 'brother' > A *pracar*, B *procer* (cp. Eng. *brother*).

17.8. With the dentals, things are not quite as simple. Both **t* and **dh* became *t*, but **d* became *ts*: compare PIE **kṛntom* 'hundred' > A *känt*, B *kante*; PIE **h₁rudhro-* 'red' > A *rtär*, B *ratre* 'red'; but PIE **dak-* 'bite' > B *tsäk-* 'bite'. The fate of **d* is complicated in other ways as well, chiefly by its propensity to disappear under conditions that are still obscure (e.g. **doru* 'wood' > AB *or* 'wood'; **sueh₂-ro-* 'sweet' > A *swär*, B *swäre* 'sweet').

17.9. This stop-consonant merger is most dramatic in the velars, where all nine of the PIE velar stops fell together as *k*. Compare the following nine examples of *k*, each from a different PIE velar:

**k̑* A *känt*, B *kante* 'hundred' < **kṛntom* (cp. Ved. *śatām* 'hundred')

**k* AB *luk-* 'shine' < **leuk-* (cp. Lat. *lūx*, *lūc-* 'light')

**k^w* A *ak*, B *ek* 'eye' < **h₂ek^w-* (cp. Gk. *óp-somai* 'I will see')

**g̑* AB *āk-* 'lead' < **h₂eḡ-* (cp. Lat. *agō* 'I lead')

**g* A *ok-*, B *auk-* 'increase' < **h₂eug-* (cp. Lat. *augeō* 'I increase')

**g^w* A *ko*, B *ke^w* 'cow' < **g^wou-* (cp. Eng. *cow*)

**gh* AB *ku-* 'pour' < **gheu-* (cp. Gk. *khé(w)ō* 'I pour')

**gh* A *lake*, B *leki* 'bed' < **legħ-* (cp. Gk. *lékhos* 'bed', German *liegen* 'lie')

**g^wh* AB *tsäk-* 'burn' < **dheg^wh-* (cp. Ved. *dahanti* 'they burn')

The labiovelars may have still been labiovelars in Proto-Tocharian. Some clear examples where a labial element was still present, or rounded a following vowel, include A *kus*, B *k_use* 'who?' < **k^wis*, and A *kukäl*, B *kokale* 'chariot' < **k^wek^wlo-* 'wheel' (cp. Gk. *kúklos* 'wheel'). The symbol *k_u* in *k_use* reflects a peculiarity of the Tocharian writing system in spelling this word: the vowel *u* was written with a subscript sign, as opposed to the other vowels which were written superscript. Tocharian B sometimes even has the sequence *kw* for a labiovelar or combination of velar plus **u*, as in *yakwe* 'horse' (A *yuk*) < **ék_uos*.

17.10. Palatalization. The remaining obstruents found in the Tocharian sound inventory are the result of palatalization before **e* or **i* in the prehistoric period; these two vowels later changed in various ways (see below), with the result that a palatalized consonant is often our only evidence of the erstwhile presence of an **e*, **i*, or the glide **j*. The palatalization rules are many and complex, and will not be treated here in their entirety. The basic outlines are that *t* usually became the affricate *c* (phonetically [č]) before *e* or *i*, as in B *pācer* 'father', and *k* became the sibilant *ś*, a *sh*-like sound, as in A *śpāl* 'head' (< **ghebh_h-(e)l-*, cp. Gk. *kephalē*). There were additional changes in consonant clusters; for example, the palatalization of **st* was *śś* or *śc*, as in A *kaśśi*, B *keścye* 'hungry', the derived adjective from A *kaśt*, B *kest* 'hunger'. Certain resonants were palatalized (see the next section below), and the glide **y* could be palatalized to *y* in Tocharian B, as in *yente* 'wind' < **h₂ueh_hnto-* (cp. Eng. *wind*).

17.11. Resonants. The resonants remained intact: A *lāc* 'he went out' < **h₁ludh_h-et*; B *mācer* 'mother' < **mātēr*; A *want* 'wind' < **h₂ueh_hnto-*. The resonant **r* was the only consonant that was preserved word-finally. The liquid *l* and the nasal *n* could be palatalized to *ly* and *ñ*, as in A *klyu* 'fame' < **k₁leu-* and AB *ñu* 'nine' < **neyn_h*.

A word on the Tocharian nasals is in order here. Besides *m*, *n*, and *ñ*, the language had a velar nasal *ŋ* (pronounced [ŋ] as in *sing*) and a letter transliterated as *ŋ* which is used to write nasals in word-final position. In the Brāhmī script this normally indicates nasalization of the preceding vowel, but in Tocharian it seems to have stood for *n*. A final *-*m* became *-*n* (as also in Greek and Celtic), which we know from the oblique stem of the word for 'earth', A *tkan-*, B *ken-* (nominative A *tkam₁*, B *ken₁*), ultimately from PIE **dh(e)ǵhōm* (cp. §6.30). (Note incidentally that A *tkam₁*, like its cognate Hitt. *tēkan*, preserves the original dental-velar order of the "thorn" cluster; recall §3.25.)

The syllabic resonants developed a prothetic **ā* in Proto-Tocharian (see §17.14 on this vowel): PIE **k₁mtom* 'hundred' > A *kānt*, B *kante*; **dn̥ǵhu-* 'tongue' > A *kāntu*, B *kantwo* (with reversal of the order of the two stops); **bh₁ǵh-ro-* 'high' > A *pārkar*. As in Germanic, **RHC* sequences normally lost the laryngeal and developed the same as ordinary **RC* sequences, as in B *pāllent* 'full' (of the moon) < **p₁h₁no₁uent-* and B *pārweṣṣe* 'first' < **p₁h₂uo-*.

17.12. Laryngeals. The laryngeals, when vocalized between consonants, became **a* in Proto-Tocharian, as in several other branches (Italic, Celtic, Germanic, Armenian); this then became *ā* in both A and B (see §17.17 below). Thus PIE **ph₂tēr* 'father' became A *pācar*, B *pācer*. Interestingly, at least some of the time the sequence **ih₂* was syllabified as **i_h₂* and became **ya*, just as in Greek. Thus the feminine accusative

singular **-ih₂-m* became **-yam* and ultimately *-ā* (preceded by a palatalized consonant) in the feminine oblique ending *-ām* (with an added particle *-m*), as in A *pontsā-m* 'all' < **pant-ih₂m(-)* (cp. Gk. *pánt-* 'all'). The same was true sometimes of **uh₂* or **uh₃*, as in B *lwāsa* 'animals' from **luHs-* 'louse'. At other times, however, **iH* and **uH* became **ī* and **ū*, whose fates are treated below in §17.18.

VOWELS

17.13. The development of the PIE vowels in Tocharian is very complicated, and unusual from the point of view of the other older IE languages in many respects. Not all of the details are agreed upon, but it is clear that the two Tocharian languages each went their own way in their development of the vowel system after the Common Tocharian period.

17.14. The vowel inventory. Both Tocharian languages possessed the two high vowels *i* and *u*, the mid vowels *e* and *o*, and three vowels transliterated as *a*, *ā*, and *ä*. The last of these, called the *Fremdvokal* or 'foreign vowel', is of disputed phonetic value. In the Sanskrit Devanāgarī script, derived from the same source as the Tocharian script, the vowel *a* represents schwa [ə], while *ā* represents a true long [a:], not only longer but also lower than a schwa. The Tocharian vowel *ä* is written in the Brāhmī script as an ordinary *a* with two dots over it; among other things, it is the usual epenthetic vowel in Tocharian, used for breaking up consonant clusters. The cross-linguistically most common epenthetic vowel is schwa, but on the assumption that the ordinary *a* represents schwa as just stated, the *Fremdvokal* must represent a different sound. It has been proposed that it represents the mid high vowel [ɨ] (found in some pronunciations of the second vowel of English *singin'* in relaxed speech).

17.15. Vowel mergers. Proto-Tocharian underwent three important mergers of the inherited vowel inventory. PIE **e*, **i*, and **u* all fell together as **ā* in Proto-Tocharian (but only after **e* and **i* had palatalized preceding consonants as per §17.10 above). This Proto-Tocharian **ā* became variously *a*, *ä*, or *ā*, depending on conditioning factors we need not go into: A *śām* and B *šana* 'wife' < **g^wen-* 'woman' (with palatalization of the old **k* to *ś*); A *wās*, B *wase* 'poison' < **yis-*; A *rtār* (< earlier **rātār*) and B *ratre* 'red' < **h₁rudbro-*. It could also be rounded to *u* when next to an old labiovelar, as in A *kumseñc* 'they come', with *kum-* < **k^wām-* < PIE **g^wm-*.

17.16. The second vowel merger was that of PIE **o* and **ē*, which became a mid front vowel notated variously as **e* or **æ* in Proto-Tocharian; this developed to *a* in A and *e* in B. Thus **ok^w-* (earlier **h₂ek^w-*) became A *ak* and B *ek* 'eye', and **ph₂tēr* 'father' became A *pācar* and B *pācer*. Interestingly, *o* remained *o* if the following syllable contained *u*, as in AB *or* 'wood', apparently from **doru* with loss of the **d* (§17.8).

17.17. The third vowel merger is that of PIE **a* and **ō*, which, together with the vocalized laryngeals (§17.12 above), became Proto-Tocharian **a*. This then became *ā* in both A and B: AB *āk-* 'lead' (< **aġ-* < **h₂eġ-*); AB *knā-* 'know' (< **ġnō-* < **ġneh₃-*, cp. Eng. *know*). This development was not shared by PIE **ā*, which became a vowel symbolized variously as **o* or **ā* in Proto-Tocharian; this turned into *a* in A and *o* in B, as in the word for 'brother' (**bhrātēr*), A *pracar* and B *procer*.

17.18. After all of these somewhat confusing developments, it will be a relief to know that PIE **ī* and **ū* (from **iH* and **uH*) simply became *i* and *u* in both

languages: A *wiki* and B *ikām* ‘twenty’ < **uikm̐ti* (cp. Lat. *uīginti*), B *suwo* ‘pig’ < **sū-* (cp. Lat. *sūs*, Eng. *sow*).

17.19. **Diphthongs.** The diphthongs beginning with *a* and *o* are reasonably well preserved in Tocharian B, but became monophthongs in A: B *aikare*, A *ekār* ‘empty’ < **aigro-* (cp. Lat. *aeger* ‘poor’); B *auk-* and A *ok-* ‘increase’ < **aug-* (earlier **h₂eug-*); and B *aise* ‘force’ < **oiso-*.

17.20. **Syncope.** The Fremdvokal *ä* disappeared in unstressed open syllables in B, and in essentially any open syllable in A. This created a number of odd word-initial consonant clusters, as in B *yšiye*, A *wše* ‘night’, AB *lkaisi* ‘to see’, B *wtentse* ‘for a second time’, A *pkänt* ‘without’.

Final syllables

17.21. Final syllables were generally eliminated in Tocharian A but preserved in Tocharian B. Their elimination in A created some awkward consonant clusters at the ends of words which were then broken up by the insertion of an *ä*, as in the word for ‘red’, *rtär*, where originally there was no vowel between the *t* and the *r* (cp. B *ratre*; PIE **h₁rudhro-*).

Stress

17.22. In Tocharian B, pairs such as B *kante* ‘hundred’ ~ pl. *käntenma*, and *äke* ‘end’ ~ pl. *akenta*, where there are alternations between *a* and *ä* and between *ā* and *a*, suggest that there was a shift of the stress from the first to the second syllable in words of more than two syllables (thus *kānte* ~ *käntenma*, *āke* ~ *akenta*). There are similar alternations in B between *a* and zero, stemming from an original stressed **ā* and unstressed **ä* (cp. §17.20 above), as in *camel* ‘birth’, pl. *cmela*, and *yapoy* ‘land’, pl. *ypauna*.

Morphology

NOUNS

17.23. **Cases.** Tocharian noun inflection differs significantly from that of almost all the other IE languages by actually having *more* cases than what we reconstruct for PIE. Interestingly, this did not come about by expanding the original PIE set; rather, Tocharian first drastically reduced the inherited case inventory and only later developed new case-endings of its own design. The PIE dative, instrumental, ablative, and locative were all lost; only four of the original cases survived into Proto-Tocharian, the nominative, vocative, genitive, and accusative. The accusative is continued as a case called the oblique (although not all obliques come from old accusatives), and the vocative only lives on in Tocharian B. These are often called the *primary cases*.

To express the grammatical relationships that had been expressed by the lost cases, Proto-Tocharian made use of postpositions (much as English uses prepositions for the same purpose). Over time, many of these were reanalyzed as case-endings, and at the end of the day the Tocharian languages had developed seven new cases, the so-called *secondary cases*. Although the case systems are nearly identical in the two languages and came about in the same way, most of the actual endings are not common to both languages – evidently, each language reanalyzed different post-

positions. One, perhaps two, are cognate: the new locative and perhaps the perlativ (expressing 'by' or 'through' someone's agency). Three others, a comitative (expressing accompaniment), allative (place towards which), and a new ablative, are found in both languages but with different endings. In addition, Tocharian A has a new instrumental case, and Tocharian B has a causal, whose function is essentially the same as the instrumental in A. The new case-endings are mostly added to the oblique, presumably once the case governed by the erstwhile postpositions. If this case system evolved under the influence of Uralic or Turkic languages, as has been suggested (see §17.4 above), the influence must have come fairly late, given that each Tocharian language developed most of the new system on its own and the endings do not usually show the expected vowel-weakenings. But the use of postpositions closely associated with their nouns was already a Common Tocharian feature.

To give an idea of the case system, paradigms of the word for 'father' in B and 'woman' in A follow. Not all the case-forms below are actually attested for these words. The vocative in B for this declension is the same as the nominative.

	B		A	
	sing.	pl.	sing.	pl.
Nom.	<i>pācer</i>	<i>pācera</i>	<i>k_uli</i>	<i>k_ulewāñ</i>
Obl.	<i>pātār</i>	<i>pācera</i>	<i>k_ule</i>	<i>k_ulewās</i>
Gen.	<i>pātri</i>	<i>paceramts</i>	<i>k_uleyis</i>	<i>k_ulewāssi</i>
Instr.	—	—	<i>k_uleyo</i>	<i>k_ulewāsyo</i>
Perl.	<i>pātārsa</i>	<i>pacerasa</i>	<i>k_uleyā</i>	<i>k_ulewāsā</i>
Com.	<i>pātāmpa</i>	<i>pācerampa</i>	<i>k_uleyaššāl</i>	<i>k_ulewāšaššāl</i>
All.	<i>pātārś(c)</i>	<i>pāceras(c)</i>	<i>k_uleyac</i>	<i>k_ulewāsac</i>
Abl.	<i>pātārmem</i>	<i>pāceramem</i>	<i>k_uleyāš</i>	<i>k_ulewāsāš</i>
Loc.	<i>pātārne</i>	<i>pācerane</i>	<i>k_uleyam</i>	<i>k_ulewāsam</i>
Caus.	<i>pātārñ</i>	<i>pācerañ</i>	—	—

17.24. "Gruppenflexion." Another feature of Tocharian that is unusual from the point of view of the rest of Indo-European is the phenomenon whereby phrases in a secondary case often exhibit a secondary case-ending on the last word only; the other words are in the oblique. This is known by its German name, *Gruppenflexion* or 'group inflection'. Thus B *kektseñ* (obl.) *reki* (obl.) *palskosa* (perlativ) 'with body, word, (and) thought' ('thought' is perlativ); A *yātālwātses* (obl. pl.) *tsopats-tampes* (obl. pl.) *nermīšmās* (obl. pl.) *wrašaššāl* (comitative pl.) 'with the powerful, mighty, artificial beings' ('beings' is comitative).

17.25. Number. Tocharian has also innovated by adding to the inherited singular, dual, and plural numbers a paral, a kind of dual used only for naturally occurring pairs such as hands, eyes, etc., such as B *ešane*, A *ašam* 'both eyes'. It was formed by suffixing a particle *-nō to the inherited dual. Tocharian B also has a "plurative" in *-auwenta* (the plural of the old word for 'one', *oi-wo-), to denote 'one at a time, individually'.

17.26. Gender. The three-gender distinction of PIE is kept intact in Tocharian, although the neuter is a living category only in the pronouns. In nouns, the descendant of the PIE neuter has masculine endings in the singular and feminine endings in the plural.

Tocharian A has the very unusual feature of distinguishing gender in the singular of the first personal pronoun 'I': masculine nominative *nās*, feminine *ñuk*. It has been plausibly suggested that the feminine form ultimately descends from the PIE nominative singular **eǵoh₂*, while the masculine continues the accusative **me*.

VERBS

17.27. The Tocharian verbal system is built around a fundamental opposition between the ordinary verb and its associated causative, which in the present tense is usually formed with a suffix going back to PIE **-skē/o-* (§5.31), a suffix which does not have causative value in the rest of IE except for a few verbs in Greek. Not all verbs have both a base form and a causative form. The causative does not always appear to have a different meaning from the base, as in B *talāššām* 'he raises', the "causative" of the synonymous verb *tallām* 'he raises'. Sometimes the base verb is intransitive while the causative is transitive, e.g. B *tsālpētār* 'is redeemed', causative *tsālpāššām* 'redeems'. In a third group of verbs, the causative has real causative value: B *kārsanam* 'he knows', causative *šarsāššām* 'he causes to know, informs'.

17.28. Verb stems. A Tocharian verb has three stems: present, preterite, and subjunctive. The present stem is used to form the present tense, the imperfect, and the present participle. The imperfect is formed differently in the two languages. In B, it is characterized by the stem-vowel *-i-* (e.g. *klyaušim* 'I heard', imperfect of *klyaušām* 'he hears'), which comes from PIE **-ih₁-*, the zero-grade of the optative morpheme. For the semantic development, compare Eng. *he would go* in the meaning "he used to go, was going." In A, the imperfect stem vowel is *-ā-*, whose origin is uncertain.

The Tocharian present stems are divided into twelve classes, which together continue most of the PIE types of present stems, including root athematic presents (Class I, e.g. A *swiñc* 'they rain' [said of flowers] < **sub₂-énti*), nasal-infix presents in B (Class VII, e.g. *pinkem* 'they paint' < **pi-n-g-*, cp. Lat. *pingunt* 'they paint'; also Class VI, formed with the suffix *-nā-* that comes originally from **-n-H-*, that is, a nasal infix to a laryngeal-final root; e.g. AB *musnātār* 'lifts up', cp. Ved. *muṣṇāti* 'steals' and §10.42), thematic verbs (Class II, e.g. B *akem* 'we lead' < **h₂eg-o-mes*, cp. Lat. *agimus* 'we drive, lead'), denominative verbs (Class XII, e.g. B *lareññentār* 'they love' < Proto-Toch. **lāren-yā-*, cp. B *lareñ* (pl.) 'dear'), and **-skē/o-* verbs (most clearly Class IX in B, e.g. *aiskau* 'I give'). Tocharian also has numerous examples of presents in **-se/o-*, which is not a common stem-formant elsewhere in IE; these form Class VIII and are especially well represented in A (e.g. *nāmseñc* 'they bow down to, revere', cp. Gk. *nēmō* 'I pay, deal out').

Of great interest for the history of PIE verbal morphology are Classes III and IV, which contain mostly middle verbs having **-o-* as the stem vowel throughout, such as B *lipetār* 'is left over' (Class III) < **lip-o-tor* and B *osotār* 'dries' (Class IV) < **as-o-*. The stem-vowel **-o-* is originally the archaic 3rd sing. middle ending **-o* (see §5.15), which later got generalized throughout the paradigm as the thematic vowel and to which new middle endings were added.

17.29. The subjunctive stem is used to form the subjunctive and optative. The subjunctive doubles as a future tense (an inherited feature; see §5.52), while the optative is used in contrary-to-fact statements. The fact that the subjunctive stem is different from the present stem is an interesting feature shared with the *ā*-subjunctive of Italic and Celtic (though not in all details); compare Lat. present indicative *attig-it* 'touches', (archaic) subjunctive *attig-at* '(that) he touch', or OIr. present indicative

-*cren* 'buys', subjunctive -*cria*. A Tocharian example is B present *kārnāstār* 'buys' (< **k^wrinh₂-sĕ-*), subjunctive *kārnātār* 'he will buy' (< **k^wrinh₂-*).

17.30. From the preterite stem are formed the preterite tense and a preterite participle. The Tocharian preterite mostly continues the PIE aorist, but a few preterite participles are from perfects.

17.31. Personal endings. The personal endings are largely the familiar ones inherited from PIE, though some of the details are not fully clear. The active and middle endings of the present tense can be illustrated from both languages with the following paradigms of B *klyausau* and A *khyosam* 'I hear':

	B	A
Active	sg. 1 <i>klyausau</i>	<i>khyosam</i>
	2 <i>klyaušt(o)</i>	<i>khyošt</i>
	3 <i>klyaušäm</i>	<i>khyošäs</i> (?)
	pl. 1 <i>klyausem(o)</i>	<i>khyosamäs</i>
	2 <i>klyaušcer</i>	[<i>khyosac</i>]
	3 <i>klyausem</i>	<i>khyoseñc</i>
Middle	sg. 1 <i>klyausemar</i>	<i>khyosmār</i>
	2 <i>klyauštar</i>	<i>khyoštār</i>
	3 <i>klyauštār</i>	<i>khyoštār</i>
	pl. 1 <i>klyausemt(t)ār</i>	<i>khyosamtār</i>
	2 <i>klyauštār</i>	—
	3 <i>klyausentār</i>	<i>khyosantār</i>

The PIE personal endings shine through clearly in some of the forms, though by no means all. The 1st sing. active B -*au* is from *-*o-mi*, that is, the thematic vowel plus the primary 1st singular ending (the *-*m-* was weakened regularly to -*w-*). The 2nd sing. active -*št(o)* is reminiscent of forms like German *bist* '(thou) art', where the -*t* comes ultimately from the 2nd sing. pronoun **tu* added on to the ending. Both 3rd singular active endings are problematic. As for the middle voice, note that Tocharian has generalized *-*r* as the middle marker, just like Italic and Celtic. The difference between 3rd pl. active A -*ñc* and B -*m* is thought by some to reflect a difference between PIE primary *-*nti* and secondary *-*nt*, but this need not be so (there is otherwise no pattern of the A forms coming from primary verbal endings and the B forms coming from secondary endings; the 1st sing. -*au* in B is from primary *-*o-mi*, for instance).

Tocharian A text sample

17.32. The poetic verses beginning the Buddhist tale known as the Puṇyavantajātaka. Each line contains 14 syllables divisible into two half-lines of seven syllables each. The last stanza consists of only one line.

- 1 *kāsu ñom-klyu tsrašišī śāk kālymentwam sātkatār.*
yārḱ yñāñmune nam poto tsraššuneyā p_ukāš kālpnāl;
yuknāl ymārāk yasluñcās, kālpnāl ymārāk yātlune.

- 2 tsrašiši māk nišpalntu, tsrašiši māk škam šñaššēn.
nāmseñc yāsluš tsrašisac, kumseñc yārkant tsrašisac.
tsrašiñ waste wrasašši, tsrašiši mā praski naš.
- 3 tämyo kāsū tsraššune p_ukam pruccamo ñi pālskam.
- 1 Good fame of the strong spreads out in ten directions.
(One) must achieve honor, worth, reverence, (and) flattery from everyone through strength.
(One) must quickly conquer enemies; (one) must quickly gain ability.
- 2 The strong have many riches, the strong also have many relatives.
Enemies bow down to the strong, honors come to the strong.
The strong (are) the protection of creatures; the strong have no fear.
- 3 Thus strength (is) good, best of all in my opinion.

17.32a. Notes. 1. *kāsu*: 'good'. *ñom-klyu*: 'fame', literally 'name-fame'. The corresponding word for 'name' in B is *ñem*; the two forms together must go back to a PIE **h₂nēh₃m₂* with lengthened-grade. Toch. A *klyu* and B *kālywe* continue PIE **klēyos*, cp. Gk. *klé(w)os* 'fame'. For the phrase *ñom-klyu* compare Gk. *onomā-klutos* 'famous for his name'; the collocation of these two roots goes back to PIE (§2.47). *tsrašiši*: 'of the strong', genit. pl. of *tsraši*. The source of the genit. pl. ending *-šši* is unknown. *šāk*: 'ten', B *šak*, from PIE **dek_m*, with palatalization of initial dental and loss of final nasal. *kālymentwam*: 'directions', loc. pl. of *kālyme*. The ten directions are the four cardinal directions, the four directions in between, plus up and down. *sātkatār*: 'spreads out', 3rd sing. Class III middle present; B *sātketār*, with *-e-* from PIE **-o-* (see §17.28). There is no convincing etymology, but the verb belongs to a group whose base ends in *-tk-* that presumably continues **ske/o-* verbs added to a root-final dental. *yārk*: 'honor', B *yarke*; PIE **erk^hos*, also in Ved. *ārcati* 'praises' and *ṛk* 'song of praise' (in *Ṛg-veda*, the Rig Veda). *ynāmūne*: 'worth', a noun derived from the adjective *ynāñm* 'worth(y)'. *nam*: 'reverence', perhaps borrowed from Skt. *namas* or perhaps a native inheritance. *poto*: 'flattery'; PIE root **bheu^hdh-* 'be aware', with derivatives meaning 'make aware, give notice'. Compare the Avestan phrase *namō baoδaitiēiti* 'bids reverence', with the same two roots collocated as in *namoto*. *tsraššuneyā*: 'with strength', perlativ of *tsraššune*. *p_ukāš*: 'from everyone', abl. of *puk* 'every, all'. *kālpnāk*: 'must achieve', gerundive, expressing necessity. The subject is unexpressed. *yuknāk*: 'must conquer', gerundive. *ymārāk*: 'quickly'. *yāsluñcāš*: 'enemies', oblique pl. of *yāslu*; the nomin. pl. *yāsluš* is in the next stanza. *yātluñe*: 'ability', abstract verbal noun from *yāt-* 'be capable', cognate with Ved. *yātate* 'is in the right place'.

2. *māk*: 'much, many'. *nišpalntu*: 'riches', pl. of *nišpal* 'wealth'. *škam*: 'also, and', a post-positive (enclitic) conjunction. *šñaššēn*: 'relatives', a derivative of *šñu* '(one's) own'. *nāmseñc*: 'they bow down', 3rd pl. Class VIII present; see §17.28. *tsrašisac*: 'to the strong', allative pl. *kumseñc*: 'they come', 3rd pl. Class X present; *kum-* continues PIE **g^hem-*, zero-grade of **g^hem-* 'come, go' (Lat. *ven-iō* 'I come', Eng. *come*); see §17.15. *yārkant*: 'honors', pl. of *yārk* above. *tsrašiñ*: 'the strong', nomin. pl. The repetition of this adjective in this stanza in different case-forms is a stylistic figure of IE poetry known as polyptoton. Note that the first two occurrences begin the first two half-lines, the next two end the next two half-lines, and the final two begin the last two half-lines of the stanza. *waste*: 'protection'. *wrasašši*: 'people', genit. pl. of *urasom*. The paradigm is irregular, with the nominative, oblique, and perlativ sing. formed to the stem *wras(o)m-*, and the other cases to the stem *wras-*. *mā*: 'not'; PIE **mē*, which in the other daughter languages is usually the negative reserved for prohibitions, but in both Tocharian languages becomes the ordinary negative (A has another negative, *mar*, used with prohibitions). *praski*: 'fear', B *proskiye*; one of a class of nouns (another being A *uše*, B *yšiye* 'night') whose suffix comes from PIE **-ōj*, very rare elsewhere in IE. The noun is ultimately cognate with Eng. *fright*. *naš*: 'there is', 3rd sing. Class II present of *nas-* 'be', PIE **nes-* 'return home, be at home'; *naš* appears to continue an unexpressed *o*-grade

**nos-ti*. (Compare perhaps the unusual *o*-grade of Gk. *nóstos* ‘homecoming’, as discussed in §12.65a.)

3. *tämyo*: ‘therefore, thus’, derived ultimately from the PIE demonstrative stem **to-*. *p̄kam* *pruccamo*: ‘the best of all’; *p̄kam* is locative of *puk*, and *pruccamo* means ‘great, wonderful’. As in Hittite and various archaic usages elsewhere in IE, Tocharian did not overtly mark the comparative or superlative degree. *ñi*: ‘of me’, genit. of *nāš* ‘I’. *pālskam*: ‘opinion’, locative of *pālsāk*.

Tocharian B text sample

17.33. From Text 18 in Wolfgang Krause and Werner Thomas’s *Tocharisches Elementarbuch* (Heidelberg, 1960–4); one of the MQ texts (see §17.2). The sounds *a*, *ā*, and *ä* are not always distinguished from one another in this text.

snai preñke takoy sa kenä yke postām po wars-ite, ešnesa menkitse takoy kacāp ompā pärkre šäyeñca, pyorye šāp takoy cew warne somo lyautai länktsa mā klyeñca kante pikwala epin̄kte kaccap su no tālašši ašco, rāmoytār rmer ka, cpi ašce lyautaiyne tā_u salkoytār kewcā: tusa amāskai lwasāmep onolmeṛtsā yšamna cmetši

Suppose that the earth were continually and entirely full of water without an island, and suppose there were a tortoise there living a long time, lacking eyes; suppose furthermore that on that water there were also a yoke with one hole, buoyant and not standing still; if in the course of 100 years this tortoise should now lift up his head, and indeed should quickly bend it, and if it should happen that his head were raised up high through that hole, then this (= the following) [would be] more difficult – for beings to be [re]born from animals among (= as) men.

17.33a. Notes (selective). *snai*: ‘without’, cp. Lat. *sine* ‘without’. *preñke*: ‘island’. *takoy*: ‘(as though it) were’, 3rd sing. optative of the verb ‘be’; usually spelled *tākoy* (as below). The Tocharian paradigm for ‘be’ is mostly constructed from two roots, *nes-* (B *nas-*; PIE **nes-*) and *tā-* (PIE **steh₂-* ‘stand’; compare Spanish *estar* ‘be’ from Lat. *stāre* ‘to stand’ for a similar semantic development). *kenä*: ‘earth, land’, PIE **dhǵhem-*. *yke postām*: ‘continually’. *po*: ‘all’. *ešnesa menkitse*: ‘without eyes’; *ešnesa* is the dual of *eš* ‘eyes’ (< PIE **h₂ekʷ-*, cp. Lat. *oc-ulus*) plus the paral ending *-ne* plus the case ending *-sa*. *kacāp*: ‘tortoise’, borrowed from Skt. *kacchapas*. *ompā*: ‘there’. *pärkre*: ‘for a long time’. *šäyeñca*: ‘living’, present participle, from PIE **ǵʰeh₂y-ont-*; the verb is cognate with Gk. *zō* ‘I live’. *pyorye*: ‘yoke’. *cew warne*: ‘in that water’; *cew* is the oblique of *su* ‘this, that’. The sequence *ew* is found only in the MQ texts. *somo lyautai*: ‘with one hole’; *lyautai* has been connected etymologically by some with Hitt. *lutai* ‘window’. *mā*: ‘nor’, the negator in Tocharian; cp. Gk. *mē*, Arm. *mi*, Ved. *mā*. *klyeñca*: ‘standing still’, present participle. *kānte*: ‘hundred’, PIE **ḱn̥tom*; normal spelling in B is *kante*. *pikwala*: ‘year’. A *p̄kāl*, B *pikul*. According to an attractive recent proposal by Joshua T. Katz, this word comes from a compound *(*e*)*pi-kʷ(e)l-o-* ‘(the thing) that turns around’, cp. Homeric Gk. *epi-tell-oménōn eniautōn* ‘as the years go turning round’. The word for ‘year’ would therefore be an example of what is known as a ‘transferred epithet’: an adjective used to describe something becomes the name of that thing, like Lat. *terra* ‘earth’ from ‘dry (land)’ (§13.31). *tālašši*: ‘raises’, PIE **telh₂-*; cp. Lat. *toll-ere* ‘to raise’. *rāmoytār*: ‘should bend’, optative of the Tocharian root *rām-*. *rmer*: ‘quickly’; also spelled *ramer* in B (A has *ymār*). It may be related to Gk. *é-dramon* ‘I ran’. *cpi*: ‘his’, also spelled *cwi*. *salkoytār*: ‘should be raised’. *kewcā*: ‘high’; the MQ spelling of B *kauc*, A *koc*. *amāskai*: ‘(more) difficult’; not morphologically a comparative (Armenian also uses the

positive degree to express the comparative). *Iwasāmēṃ*: ‘from animals’; the ‘animal’ word is *A lu*, *B luwo*. It has been compared with the word for ‘louse’ in other IE languages, probably from PIE **luHs-*. *onolmentsā*: ‘for beings’, from *onolme* ‘a being’, Proto-Toch. **ān-elme*, from **ān-* ‘breathe’ (PIE **h₂en-*, cp. Lat. *an-ima* ‘breath’) plus a suffix. *yāmna*: ‘among men’, from **en-* ‘in’ plus the root for ‘life’, cp. *šāyeñica* above. *cmetsi*: ‘to be (re)born’, Tocharian root *tam-*.

For Further Reading

The number of book-length historical and comparative treatments of Tocharian is not extensive. A good recent overview of the historical phonology and morphology is Pinault 1989. In English, there is Adams 1988, which is somewhat difficult to use; the same author has also recently come out with an etymological dictionary of Tocharian B (Adams 1999), which partly supplants the older general Tocharian etymological dictionary that makes up most of van Windekens 1976–82. (This work also has comparative grammatical discussions and extensive references to earlier literature; but his own ideas sometimes need to be treated with caution.) The standard reference grammar is still Krause and Thomas 1960–4, the second volume of which contains texts and a glossary. The Tocharian A texts have been edited and published in Sieg and Siegling 1921, and the Tocharian B texts in Sieg and Siegling 1949–53. The latter has been partially supplanted by Thomas 1983. A journal called *Tocharian and Indo-European Studies*, which appears irregularly, has featured many excellent articles by prominent contemporary Tocharologists. The most recent technical work on historical Tocharian phonology is Ringe 1996. A recent thorough discussion of the Takla Makan mummies is Mallory and Mair 2000.

For Review

Know the meaning or significance of the following:

Tarim Basin MQ texts Fremdvokek Gruppenflexion

Exercises

1 Briefly explain the history or significance of the following Tocharian forms:

- | | | |
|------------------------------------|------------------|----------------------------------|
| a <i>A tōār</i> , <i>B tapre</i> | e <i>B yakwe</i> | i <i>B lwāsa</i> |
| b <i>B tsāk-</i> | f <i>B yente</i> | j <i>A ašām</i> , <i>B ešane</i> |
| c <i>AB or</i> | g <i>A klyu-</i> | k <i>B līpetār</i> |
| d <i>A kukāl</i> , <i>B kokale</i> | h <i>A tkam</i> | |

2 In §17.11, it was stated that Toch. *A tkam*, *B kem* ‘earth’, along with Hitt. *tekān*, preserve the original order dental–velar of the word-initial ‘thorn’ cluster; this shows that the order was reversed in *khthōn* ‘earth’ in Greek. A skeptic might say that it could have been the other way around: Greek might preserve the original order, and Tocharian and Hittite might have reversed it instead. In general terms, what evidence does the Greek verb *tiktō* ‘I give birth’ contribute to settling the argument, assuming that this is a thematic reduplicated verb from **tek-* ‘give birth’?

- 3 a PIE **dhiǵh-*, zero-grade of **dheǵh-* 'form with the hands, shape', became Toch. B *tsik-ale* 'should be made', and PIE **dhegʰh-* 'burn' became *tsak-štär* 'it burns'. Contrast this with PIE **dheh₁-* becoming Toch. A *ca-sär* 'they put', and **h₁udh-e-t* becoming Toch. B *lac* '(s)he went out', which show the ordinary development of palatalized **dh*. What must have happened to the **dh* in the two first examples?
- b What sound change, known also from Sanskrit and Greek, might have happened to affect **dhiǵh-* and **dhegʰh-* in Tocharian?
- 4 The word for 'high', A *tpär*, B *tapre* (mentioned in §17.7), has a different suffix from that seen in its Lithuanian cognate *dubùs* 'deep'. Comment on this in light of §6.87.
- 5 Given the history of the secondary cases, how would you explain the genesis of "Gruppenflexion" (§17.24)?
- 6 Briefly explain the significance of Classes III and IV of Tocharian verbs.

PIE Vocabulary IX: Form and Size

- *bherǵh-* 'high': Hitt. *parku-*, Ved. *bṛhánt-*, Arm. *barjr*, OHG *burg* 'fort'
**medh₁jo-* 'middle': Ved. *mádhyā-*, Gk. *mésos*, Lat. *medius*, Eng. MID
**meg-* 'big, great': Hitt. *mekki*, Ved. *máhi*, Gk. *méga*, Lat. *magnus*, Eng. MUCH
**mreǵhu-* 'short': Av. *mərəzu-*, Gk. *brakhús*, Lat. *brevis*, Eng. MERRY
**gʷerh₂-* 'heavy': Ved. *gurú-*, Gk. *barús*, Lat. *gravis*
**albh-* 'white': Hitt. *alpaš* 'cloud', Lat. *albus*
**h₁reudh-* 'red': Ved. *rudhirá-*, Gk. *eruthrós*, Lat. *ruber*, Toch. B *ratre*
**dens-* 'thick': Hitt. *daššu-* 'strong', Gk. *dasús*, Lat. *densus*

18 Balto-Slavic

Introduction

18.1. Balto-Slavic contains **two branches**, Baltic and Slavic (or Slavonic). The notion of a single Balto-Slavic **speech community** has been controversial in some circles, in part because of **political tensions**. But all major Indo-Europeanists are agreed that Baltic and Slavic **deserve to be grouped together**, though some dispute remains about the **exact degree and nature** of their affinity. Prehistorically, as discussed in more detail below, the **ancestors of the Balts and Slavs** were located in eastern Europe, with the Balts **probably to the northwest** of the Slavs. They were thus located near the **Germanic tribes**, and in fact there are numerous interesting features shared by Balto-Slavic and **Germanic**: both branches have dative and instrumental plurals with an ***-m-formant rather than the *-bh-formant** seen elsewhere in IE (cp. §6.17); both form a **demonstrative pronoun** with the stem ****k̑i-*** (e.g. Lith. *šis* ‘this, he’, Goth. *himma* ‘this’ [dat.]; §7.9); both have merged ***a** and ***o**; and both have a distinction between **ordinary** and **“definite”** adjectives, the latter formed with the addition of a suffix (**though not the same one in the two branches**). There are also three striking **agreements in the numeral system** not found elsewhere in older IE languages: the names for the **decads** consist of the unit plus a collective for ‘ten’ (e.g. OCS *trije desęti*, Goth. *þreis tigjus* ‘thirty’); the words for ‘thousand’ are similar (OCS *tysęštŭ*, Lith. *tūkstantis*, Goth. *þūsundi*); and the numerals eleven and twelve in Lithuanian and Germanic are expressed with phrases meaning ‘one left (over)’ and ‘two left (over)’ (Lith. *vienuolika*, *dvylika*, Goth. *ainlif*, *twalif*). It is unclear which of these features are **shared innovations**, which are areal features of north or northeast Europe, and which might be chance resemblances.

Balto-Slavic shares the **“ruki” rule** (§18.6 below) with Indo-Iranian, a sound change that is a far from trivial innovation. Some specialists have felt that this – together with the fact that both branches are satem – points to an early period of common development; others maintain that it was a feature that spread by diffusion from Indo-Iranian during a period of prehistoric contact. Arguing for the latter is the fact that **“ruki”** is far more consistently found in Slavic than in Baltic, and there is better independent evidence for early contacts between Slavic and Indo-Iranian than between the latter and Baltic.

From PIE to Balto-Slavic

Phonology overview

18.2. Balto-Slavic is defined by at least three unique phonological features in Indo-European: the development of a distinction between rising and falling pitch accents; the change of the syllabic resonants typically to resonants preceded by *i*; and the change of *VRHC to *VRC. Among the basic developments that it shares with one or more other branches, Balto-Slavic is characterized by a satem development of the velars; by the merger of aspirated and unaspirated voiced stops; by the “ruki” rule; and by the merger of **a* and **o*.

CONSONANTS

18.3. Merger of aspirated and unaspirated voiced stops. The attested Balto-Slavic languages have plain voiced stops reflecting both the PIE voiced stops and the voiced aspirates. Examples include OCS *berp* ‘I gather, take’ < PIE **bher-* ‘carry’, Russ. *gorod* ‘city’ < PIE **ghor-dho-* ‘enclosure’, and Lith. *dubùs* ‘deep’ < **dhub-u-*. It had been assumed that the two series merged by the time of Common Balto-Slavic until the Indo-Europeanist Werner Winter proposed in the 1970s that the distinction had persisted for longer, at least between **dh* and **d*. According to his original formulation, short **e* was lengthened before **d* but not **dh*, as in Lith. *sėsti* ‘to sit’ < **sed-ti* (ė writes a long *e* in Lithuanian). However, the long vowels in many of the putative examples (including this one) could also be old lengthened grades, and Winter’s Law, as it has been dubbed, remains controversial.

18.4. Velars. Balto-Slavic is a satem branch: the PIE plain velars **k* **g* **gh* and the labiovelars **kʷ* **gʷ* **ghʷ* all became the plain velars *k* and *g*: PIE **kʷos* ‘who’ > Lith. *kàs*; **sneigʷh-o-* ‘snow’ > Russ. *sneg*; **gʷerh₃-* ‘consume’ > Lith. *geriù* ‘I drink’. The palatal velars **k̑* and **g̑* became sibilants: PIE **k̑d-* ‘heart’ became OCS *sȓd-ice* and Lith. *širdis*; PIE **ǵh₁-to-* ‘golden’ became Polish *złoty* ‘gold piece, unit of currency’; and **ǵhel-to-* ‘golden’ became Lith. *želtas*.

18.5. But quite a few words do not show the expected satem development. For example, alongside Lith. *želtas* quoted above there is also *gelts* ‘yellow’ from the same PIE preform. Similarly, PIE **h₂ek-men-* ‘stone’ has *k* in most of its Balto-Slavic descendants, such as OCS *kamy* and Lith. *akmuõ*; but some Lithuanian dialects have the expected satem outcome, *ašmuõ*. From the extended root **k̑leu-s-* ‘hear’, Slavic has satem *s-* (e.g. Russian *slušať* ‘to hear’) but not Baltic (Lith. *klausyti* and Latv. *klausīt* ‘to hear’). The reverse is the case with ‘goose’, PIE **ǵhans-*: Slavic **ǵosī* (> e.g. Russ. *gus*) but Lith. *žasis*.

Various explanations have been proposed for these facts. Some of the exceptions, such as the Slavic word for ‘goose’, may be borrowings from a neighboring centum language like Germanic. But this explanation is difficult to maintain for many of these words. A more probable explanation is that there were early dialectal differences within Balto-Slavic with regards to the outcomes of the palatal velars, and that these differences have persisted. This hypothesis, unfortunately, cannot yet be tested.

18.6. The “ruki” rule. Like Indo-Iranian (§10.10), Balto-Slavic changed PIE **s* when it followed the sounds **r*, **u*, **k*, or **i*. The outcomes are different in Slavic

and Baltic, and the change was not as consistent in Baltic. See §§18.19 and 18.55 below for more details.

18.7. Resonants and glides. The resonants and glides (except **y*) remained intact: PIE **medhu-* 'mead, honey' > Lith. *medūs* 'honey'; PIE **nokʷts* 'night' > OCS *nočŭ*; **ghor-dho-* 'enclosure' > Russ. *gorod* 'city'; **leikʷ-* 'leave' > Lith. *liekù* 'I remain, let'; **iu-n-g-* 'yoke, join' > Lith. *jungti* 'to join'. The labiovelar glide **y* became *v*, as in **medh-* 'lead' > OCS *vedŭ* and Lith. *vedù*, both 'I lead'.

18.8. The syllabic resonants **r*, **l*, **m*, **n* typically became **ir*, **il*, **im*, **in*, but also sometimes **ur*, **ul*, **um*, **un* under unclear conditions. The normal outcome is preserved most clearly in Baltic, as in the following Lithuanian examples: *širdis* 'heart' < PIE **k̑rd-*; *vilkas* 'wolf' < **u̯lkʷos*; *šimtas* 'hundred' < **k̑ntom*; and *ginti* 'to drive' < **gʰm̑-tei* 'in smiting' (or the like). Note that the form *šimtas* importantly preserves the *m* before the following *t* here; in all the other IE languages it became *n* by assimilation, as in Lat. *centum*, Eng. *hund*(red), and Tocharian B *kante*.

18.9. Laryngeals. Balto-Slavic does not have any reflexes of the laryngeals per se, but it does have important indirect evidence for them in the outcomes of resonants followed by laryngeal. These sequences developed into syllables with intonational contours that differ from those of similar syllables that originally did not have a laryngeal. See below, §18.12.

VOWELS

18.10. In Balto-Slavic, the lower back vowels **a* and **o* merged, though with different outcomes in the two subbranches: in Slavic, they merged to *o*, whereas in Baltic they merged to *a*. The long vowels **ā* and **ō* also merged in Slavic (to *a*), but were kept distinct in Baltic. For examples and further discussion see §§18.24 and 18.58.

18.11. Accent and intonation. Proto-Balto-Slavic, like Vedic Sanskrit and Greek, had a mobile pitch-accent. An important innovation was the creation of intonational contrasts, in particular between a rising pitch-accent, called the *acute*, and a falling pitch-accent, called the *circumflex*. Such contrasts together with mobile accent are preserved in Lithuanian and in two South Slavic languages, Serbo-Croatian and certain Slovenian dialects; Latvian also has the pitch-accent but has fixed it on initial syllables. The other Slavic languages have replaced the pitch-accent with a stress-accent; in some it has retained its mobility (as in all of East Slavic), while in others it has become fixed or predictably placed (as in most of West and South Slavic). Those languages that preserve the mobile accent have an enormous variety of accentual rules, making historical Balto-Slavic accentology a prodigiously complicated field.

18.12. Intonation and the laryngeals. Balto-Slavic is unique in IE for turning a sequence **VRH* before consonant into **VR*. Thus for example **ġenh₁-to-* became Balto-Slavic **zēnta-* 'son-in-law', and **bherHġeh₂* became Balto-Slavic **bērza-* 'birch'. The resulting long vowels were shortened but received an acute intonation. Thus **zēnta-* became Lith. *žentas* 'son-in-law' and Serbo-Croatian *zēt*; and **berza-* 'birch' became Lith. *bērzas* 'birch' and Serbo-Croatian *brēza*. (Confusingly, the modern languages do not all use the same accent marks for writing the intonational pitches: Lithuanian *é* and Serbo-Croatian *ě* both represent *e* with original acute intonation.)

By contrast, an original short vowel plus resonant (**VR*) plus consonant received circumflex intonation. Thus PIE **ġombho-* 'tooth, toothlike projection' became Balto-Slavic **zamba-*, whence Serbo-Croatian *zŭb* 'tooth' and Lithuanian *žam̃bas* 'sharp edge', and Balto-Slavic **gelta-* or **gilta-* (also **zelta-*, **zilta-*) 'yellow' became

Serbo-Croatian *žūt* and Lithuanian *želtas* 'yellow'. (Again, the Lithuanian and Serbo-Croatian accent marks have equivalent historical value.) The different outcomes are summarized in the following table:

	PIE	Balto-Slavic	Serbo-Croatian	Lithuanian
*VRH	* <i>ǵenb₁-to-</i> * <i>bherHǵeb₂</i>	* <i>zēnta-</i> * <i>bērzā-</i>	<i>zèt</i> <i>brěza</i>	<i>žéntas</i> <i>bėrzas</i>
*VR	* <i>ǵombhos</i> * <i>ǵhel-to-</i> , * <i>ǵhl-to-</i>	* <i>zamba-</i> * <i>g/zelta-</i> , * <i>g/zilta-</i>	<i>zūb</i> <i>žūt</i>	<i>žam̃bas</i> <i>žėltas</i>

Because a syllabic resonant sprouted an *i* or *u* before it (§18.8 above), an original sequence **RHC* became **iRHC* or **uRHC*, and fell under the general category of **VRHC*, lengthening to **iRC* or **uRC* and then shortening with acute intonation. Balto-Slavic therefore distinguishes reflexes of the syllabic resonants when followed by a laryngeal from reflexes where there was no laryngeal. Compare the different accentual patterns in these descendants of the PIE words for 'full' (PIE **p̥l̥h₂no-*) and 'wolf' (**u̯lk^wos*):

	PIE	Balto-Slavic	Serbo-Croatian	Lithuanian
* <i>RH</i>	* <i>p̥l̥h₂no-</i>	* <i>p̥l̥na-</i>	<i>p̥ñ</i>	<i>p̥l̥nas</i>
* <i>R̥</i>	* <i>u̯lk^wos</i>	* <i>vilka-</i>	<i>v̥lk̃</i>	<i>vilkas</i>

(Lith. *˘* represents an original acute.) An acute accent all by itself does not automatically signal the former presence of a laryngeal, since acute intonation has other sources as well. The Balto-Slavic intonational facts therefore serve primarily for confirmation of laryngeals whose existence has been deduced from other comparative evidence.

Morphology

18.13. A characteristic innovation in Balto-Slavic morphology, found especially in verb formation, is the widespread analogical creation of so-called "neo-lengthened-grade" forms, essentially the lengthening of vowels for ablaut purposes beyond the original scope of lengthened grades in PIE. Most dramatically, inherited zero-grade forms containing the vowels **i* and **u* were often lengthened to **i̯* and **u̯* by analogy. For example, the zero-grade **m̥₁-dic* (cp. Lat. *mor-t-* 'death') became **mir-* in Balto-Slavic, and from this a "neo-lengthened" root form **mir-* was created that appears in Russ. *u-mir-at* 'to lie dying' (see §18.24 on Slavic *i* < **i̯*). Another morphological innovation of Balto-Slavic was the creation of "definite" adjectives (alongside normal adjectives), as mentioned above in §18.1. This was done by adding a suffix *-*to-*: thus OCS has *dobrŭ* 'good' alongside definite *dobrŭ-jŭ* 'the good (one)', and Lithuanian has *gėras* 'good' alongside *geràs-is* 'the good (one)'. This suffix descends from the PIE relative pronoun (recall §8.28 and compare §15.19).

In both Baltic and Slavic, the PIE verbal system was somewhat simplified. The perfect has been lost as a category, though the perfect participial suffix survives as a past participial suffix. Of some interest is the fact that both subbranches retain one tense system formed with an *s*-suffix, but a different one in each: Baltic has an

s-future lacking in Slavic (except for one isolated form, see §18.28), whereas Slavic has an s-aorist system (including secondary derivatives) lacking in Baltic.

Slavic

18.14. We know little, if anything, of the Slavs before their first mention in Byzantine histories of the sixth century AD, when bands of them moved into Greece and the Balkans from the northeast. This period also saw their expansion into parts of northern Europe previously settled by Baltic and Finnic peoples. Prior to these expansions, which continued for several hundred years and probably marked the end of Proto-Slavic linguistic unity, the Slavs are believed to have occupied an area stretching from near the western Polish border eastward to the Dnieper River in Belarus'. Archaeologically this territory shows a continuum of cultural development from about 1500 BC to the period of the expansions, which may well mean that it was Slavic (or pre-Slavic) during all that time.

18.15. Some or all of the ancestors of the Slavs were once located a little further to the east, in or near Iranian territory, for a number of Iranian words were borrowed at an early date into pre-Slavic. One of these is the word for 'god', OCS *bogŭ* (Russian *bog*), which comes from Iranian **baga-* 'god'. Two others are evidenced by Russ. *sobaka* 'dog' (cp. Median *spāka*), and perhaps by OCS *sŭto* 'hundred' (if borrowed from Iranian **sata(m)*). This territory would have been to the north of the Black Sea, an area into which a number of rivers flow whose names are usually taken to be Iranian: the Dniester, Dnieper, Donets, and Don (cp. Iranian **dānu-* 'river'). As the early Slavs moved slowly westward, they came into contact with Germanic tribes; as mentioned in §16.2, from them they borrowed the word for 'bread', OCS *xlěbŭ* (Germanic **hlaibaz*, cp. Goth. *hlaifs*, Eng. *loaf*), 'prince', OCS *kŭnєzŭ* (Gmc. **kuningaz* 'king', cp. OE *cyning*), and several others.

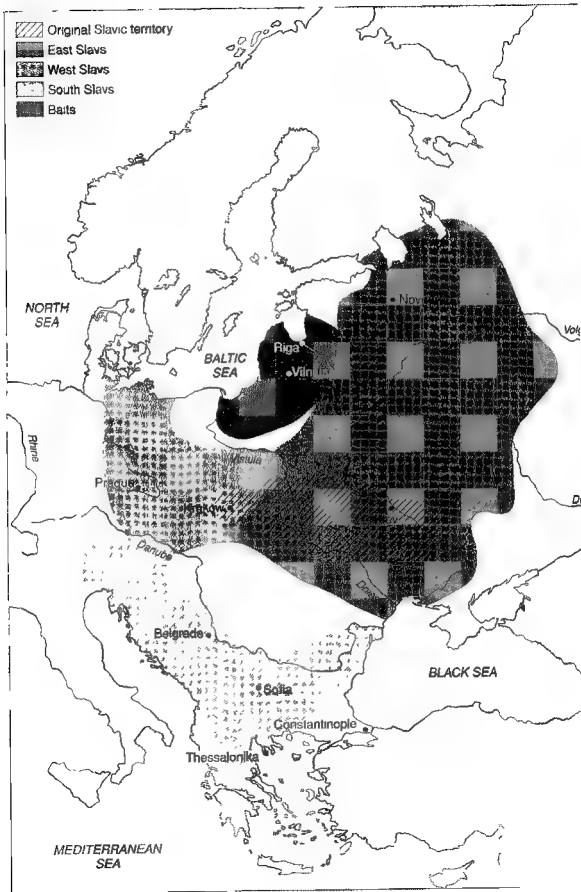
18.16. The modern Slavic languages arose as the result of dialectal differentiation that only began about 1500 years ago. Even today, several of the modern languages are to some extent mutually intelligible in their spoken forms, and even more so in their written forms. A discussion of the rise of Slavic literacy is deferred until the section on Old Church Slavonic below (§§18.33ff.), the first attested Slavic language.

Slavic phonology

18.17. Consonants. As noted above, Slavic changed PIE **k̑* and **g̑(h)* to *s* and *z*: **k̑lōyā* 'fame' > OCS *slava* 'fame, glory' (cp. Gk. *klē(w)os* 'fame' < **k̑lēyos*); **g̑neh₃*- 'know' > OCS *znati* 'know'. Aside from this change, Slavic consonants were most affected by the Law of Open Syllables and the three palatalizations discussed below.

18.18. Law of Open Syllables. A number of developments conspired to reduce the number of closed syllables in pre-Slavic, typically by deleting syllable-final consonants. This resulted in the simplification of many word-internal consonant clusters. For example, the cluster *pn* in the word **supno-* 'sleep' was syllabified *sup.no-*, with the *p* closing the first syllable; but sound changes led to the disappearance of the *p* in **sup-no-*, changing it to **su-no-* with an open first syllable. The aggregate of these changes is called the Law of Open Syllables. Two specific changes that fall under this rubric are treated below in §§18.26 and 18.27. One cluster that remained

Balto-Slavic



Map 18.1 The Slavs and Balts around AD 1000

throughout all this was **st*, seen especially in the reflex of PIE dental-plus-dental clusters (§3.36), as in OCS *vesti* 'to lead' < **ned-tei*.

As a result of the Law of Open Syllables, many words in Slavic exhibit some striking alternations, such as that between *sŭp-* and *sŭ-* in forms of the OCS verb meaning 'sleep': aorist *u-sŭp-e* 'fell asleep' but present infinitive *u-sŭ-ŋoti* 'to fall asleep' (the latter from **u-sŭp-ŋoti*; cp. Russ. *spat* 'to sleep', *usnut* 'to fall asleep').

18.19. The "ruki" rule. In Slavic, **s* occurring after **r*, **u*, **k*, or **i* became the voiceless velar fricative *x*: OCS *uxo* 'ear' < **h₂euso-*; 1st sing. aorist *rěxŭ* 'I spoke' < **rēk-s-om*; locative pl. *-ěxŭ* < **-oisu*. (The many Slavic words beginning with *x-* are unexplained, as ruki is the only known Slavic sound change that gave rise to *x*.)

18.20. Palatalizations. Of great importance in the history of Slavic consonants were a series of palatalizations; the changes are complex and will only be given here in outline. The First Palatalization caused the velars *k g x* to be palatalized to *č ž š* before **i* or a front vowel. Examples include OCS *služiti* 'to serve' < pre-Slav. **slug-iti* and *ročĭka* 'little hand, handle' < pre-Slav. **rōk-ika* (diminutive of **rōka* 'hand'). The Second and Third Palatalizations had different outcomes in different parts of Slavic-speaking territory, and happened after certain changes to the vowels had already taken place. The Second Palatalization affected the velars *k g x* when they stood before *ě* (the outcome of **ai* and **oi*, see §18.25 below): PIE **k^woi-nēh₂* 'price' > Slav. **kěna* > OCS *cěna* 'price' (*c* is [ts]; cp. Gk. *poínē* 'price, penalty', Lith. *kaina* 'price'); pre-Slav. **xairŭ* 'grey' > **xěrŭ* > Russ. *seryj*. (These are not the only outcomes; Czech, for example, has *šerý* corresponding to Russ. *seryj*.) The Third Palatalization was conditioned by the vowel preceding (rather than following) the velar: *k g x* were palatalized when followed by most vowels and preceded by *ĩ, i*, or *ę* (on these vowels, see §§18.22, 18.24, and 18.26 below); this was a late change and its effects were more sporadic. Thus the Slavic diminutive suffix **-iko-* (< PIE **-iko-*) became *-iči* or *-ice*: OCS *otiči* 'father' < **at-iko-* (**at-* also in Goth. *Att-ila* 'little father'); OCS *sřidice* 'heart' < **sřidiko-* < **křd-iko-*.

18.21. Glides. The one surviving glide, **j*, mostly stayed intact in initial position and intervocalically, as in OCS *junŭ* 'young' < **jou-no-*. Several Slavic languages have had the propensity to develop prothetic glides in front of certain vowels. For example, the number 'eight' began originally with *o-* (PIE **oktō*); in pre-Russian this developed into **yo-* and then into *vo-* (whence Russ. *vosem*, versus e.g. Polish *osiem* without the glide). Compare for this development the pronunciation of Eng. one with initial *w-* (the original glideless pronunciation is preserved in *on-ly* and *al-one*).

VOWELS

18.22. Short vowels and the "yers". As mentioned earlier, Slavic merged **a* and **o* as *o*: **nas-o-* 'nose' > OCS *nosŭ*, **h₂oyi-* 'sheep' > OCS *ovŭ-ca*. Short **e* remained *e*, e.g. **g^wen-* 'woman' > OCS *žena*. The short high vowels **i* and **u* developed into vowels transliterated as *ĩ* and *ũ* (in Cyrillic, ѣ and ѳ), called the front and back yers (or jers), respectively. These are assumed to have been very short *i* and *u*. Examples from OCS include *ěto* 'what' < **k^wi-* and *rŭžda* 'rust' < **h₁rudh-ǵā* ('the red stuff', from **h₁reudh-* 'red').

18.23. The yers are still preserved in Old Church Slavonic, but around the end of the Common Slavic period (c. AD 800–1000), they began to be lost as a category: certain ones weakened and disappeared, while others strengthened and became full

mid or low vowels. The rule was as follows: word-final yers and yers before a syllable with a full vowel (i.e., other than a yer) disappeared; these were the so-called "weak" yers. Yers before a syllable containing another yer were "strong" and later developed into full vowels. Thus in OCS *dīnīnica* 'morning star' and *rūtū* 'mouth', the first yer was strong but the second was weak; in *vūxodū* 'entrance', both were weak. The full vowels that developed from the strong yers differ from language to language; in Russian, strong *ī* and *ū* became *e* and *o*. Thus *dīnīnica* and *rūtū* correspond to Russ. *dennica* and *rot*.

The loss of the yers had far-reaching effects. It created new consonant clusters, including complex ones word-initially, as in Russ. *čtenie* 'reading' (< *čtī-*), *mglo* 'mist' (cp. OCS *mīglo*), and *vzgljad* 'view' (older *vūz-*). It also led to alternations where one form has a vowel that disappears in related forms, as in Russ. nominative *rot* 'mouth' alongside genitive *rta* 'of the mouth' (cp. the forms for 'American' in §18.31 below).

18.24. Long vowels. The long high vowels **ī* and **ū* became *i* and *y* (the latter phonetically the central high vowel [i], midway between the vowels in *pit* and *put*): OCS *u-mīr-ati* 'to die' < pre-Slavic **ou-mīr-atei* (< **mīr-*, neo-lengthened grade of **mīr-*; see §18.13); OCS *mysī* 'mouse' < **mūs-*. Long **ē* became a sound transliterated as *ě*, called *yat* (or *jat*) and representing phonetically the sound [æ] (as in *hat*), e.g. *děti* 'place, put' < **dhē-* < **dheh-*. (The apostrophe in the word *yat* represents palatalization of the final consonant in Russian, the source of the word.) Finally, long **ā* and **ō* merged as *a* (e.g. **steh-* > **stā-* > OCS *sta-ti* 'to stand', **gneh-* > **gñō-* > OCS *zna-ti* 'to know').

18.25. Diphthongs. The diphthongs were all monophthongized: **ei* became *i* (as in PIE **steigh-* 'step, go' > OCS *stignō* 'I come'); **oi* became *ě* or, word-finally, often *i* (e.g. **oisu* > OCS loc. pl. *-ěxi*, but masc. nomin. pl. **oi* > OCS *-i*); **ai* also became *ě* (e.g. **latyo-* 'left' > OCS *lěvū*); **eu* became *ju* (e.g. **bheudh-* 'be awake' > OCS *bljudō* 'I keep watch', with secondary *l*); and **ou* and **au* became *u* (e.g. **sūnou* 'son' [vocat. sing.] > OCS *synu*; **tauro-* 'bull' > OCS *turū* 'wild bull').

18.26. Rise of nasalized vowels. Common Slavic changed sequences of vowel plus nasal in closed syllables into nasalized vowels as part of the Law of Open Syllables. An original sequence **in* **en* or **im* **em* before a consonant became a front nasal vowel transliterated as *ę*: pre-Slav. **sųento-* 'holy' > OCS *svęti* 'holy'; pre-Slav. **desim-ti-* 'group of ten, decad' > OCS *desęti* 'ten'. Any other vowel plus nasal became a back nasalized vowel transliterated as *q*: **gombho-* 'tooth' > OCS *zqbū* 'tooth' (cp. Eng. *comb*), **ghans-* 'goose' > OCS *gqsī*. Later, in most Slavic languages – the main exception being Polish – these were denasalized, whence e.g. Russian *svjatoj*, *desjat*, *zub*, and *gus* for the words above (but Polish *święty*, *dziesięć*, *zab*, *gęś*).

18.27. "ToRT" groups. One who is familiar with a bit of Russian geography may recall that some cities have names ending in *-grad* (e.g. *Lenin-grad*, *Kalinin-grad*), while others have names ending in *-gorod* (e.g. *Nov-gorod*). Both elements mean 'city', but one is the native Russian word (*gorod*) while the other is borrowed from OCS (*gradū*). Russian has many other such doublets, including *golova* 'head' vs. *glava* 'head (= leader), chief' and *moroka* 'twilight' vs. *mrak* 'gloom'. These doublets exemplify different dialectal Slavic outcomes of the Proto-Slavic sequence **oR* between two consonants, conventionally labeled "ToRT"; this sequence underwent various reconfigurations as part of the Law of Open Syllables. In South Slavic, this sequence became *Ra*, while in East Slavic the outcome was *oRo*. Thus Common

Slavic **gord-* **golv-* **mork-* are the ancestors of OCS *gradŭ glava mrakŭ* and Russian *gorod golova morok-*. (OCS vocabulary is mined for scientific or learned words in Russian, much as Latin and Greek are mined by English for similar purposes.)

Morphology

VERBS

18.28. Slavic continues the PIE present, imperfect, and *s*-aorist. Only one *s*-future form is found, Old Russ. (Russ. Church Slavonic) *byšęštŭ*, a participle meaning 'about to be' (< Balto-Slavic **bŭ-sint-ŷ*; cp. §5.36). Several different participles are found, most of them continuations of IE formations: a present active participle in *-ęt-* and *-ęt-* from **-elont-*; a present passive in **-mo-* (see §5.56); two past active participles, one in *-vŭš-* (from the PIE perfect participle) and one (called the resultative participle) in *-l-* (see §5.56); and two past passive participles, in *-t-* and *-(e)n-*, from the old verbal adjectives in **-tō-* and **-nō-* (§5.57). Note that Baltic does not have the *l-* or the *(e)n-* participles. The Slavic infinitive in *-ti* (< **-tei-*) is an old case-form, probably a locative, of a verbal abstract noun in **-ti-* (§5.54).

18.29. **Perfective and imperfective aspect.** The singular hallmark of the Slavic verb is its aspectual system. Almost every verb consists really of a pair of verbs, one used to express imperfective aspect (action that is incomplete, ongoing, repeated, or unspecified as to completion), and one to express perfective aspect (completed or one-time action). Perfectives often continue present-tense formations in PIE, which then acquired future meaning in Slavic; thus for example PIE (unreduplicated) **deh₃-mi* 'I give' became the perfective OCS *damŭ* 'I will give'. However, not all PIE presents underwent this change; thus **bheroh₂* 'I carry' (extended in pre-Slavic to **bherōmi*) became the imperfective present *berę* 'I gather, take'. Very often a preverb marks the perfective member of the pair, as in Russ. *pro-čitaju* 'I will read (and finish reading)', perfective of *čitaju* 'I am reading'; here English offers a parallel with such pairs as *drink* vs. *drink up*, *eat* vs. *eat up*, etc. Many Slavic perfectives are nasal presents, e.g. OCS *stanę* 'I will stand [up]', Russ. *vernus'* 'I will return'.

NOUNS

18.30. The rich nominal inflectional system of PIE is practically intact in Slavic. Seven cases are expressed: nominative, genitive, dative, accusative, instrumental, locative, and vocative. The genitive of *o*-stems in Slavic (and Baltic) is formally the old IE ablative in **-ōt*, which took over the functions of the genitive and replaced it; this became *-a* by the change of **ō > a* and loss of the final stop. The functional merger of ablative and genitive is why prepositions with the meaning 'away from' take the genitive case in Slavic. (Compare Greek, where the ablative disappeared and its functions were taken over by the genitive, with the same results.) In feminine nouns, the old accusative plural has also taken over the nominative plural. Dual forms were still in use in the early periods, but have disappeared from most of the modern languages except dialectal Sorbian and Slovenian.

18.31. An unusual morphosyntactic innovation of Slavic is the distinction between masculine animate and inanimate accusatives. An inanimate masculine accusative is not distinguished from the nominative (thus Russ. *kabinet* 'office' is both nominative and accusative singular, and *kabiny* is both nominative and accusative plural), but

the animate masculine accusative has the same ending as the genitive (thus Russ. *amerikanec* 'American' is nominative singular, while *amerikanca* is both accusative and genitive singular; *amerikancy* 'Americans' is nominative plural, but *amerikancev* is both accusative and genitive plural).

NUMERALS

18.32. The numerals occupy a rather interesting corner of Slavic morphology. The system of numerals in almost any Slavic language is by English standards enormously complicated. In Russian, for example, the number 'one' and any compound number ending in 'one', like 21, 451 (but not 11), is an adjective agreeing with the following noun in gender and case: *odin stol* 'one table', *odno okno* 'one window', instrumental *s odnim stolom* 'along with one table'. The numbers 'two', 'three', and 'four', as well as compound numbers ending in these (such as 52, 73, 164, but not 12–14!), take the genitive singular of whatever noun follows them (e.g. *tri stola* 'three tables', where *stola* is genitive singular), and all other numbers – five through twenty and any compound number ending in five through zero – take the genitive plural of the following noun (e.g. *devjat' stolov* 'nine tables', with *stolov* genitive plural).

The historical explanation for this is an interesting example of the interrelationship of morphology and syntax. First, Common Slavic revamped the inherited cardinal numerals from five through ten by replacing them with collectives. These collectives were abstract *i*-stem nouns (e.g. OCS *peti* 'five' < **penk**(e)-ti-, *osmi* 'eight' < **oktyn*-mi), and due to their literal meaning ('a group of five, a quintet', etc.) took the genitive plural of whatever noun followed. Thus OCS *peti stolū* 'five thrones' historically meant 'a quintet/fivesome of thrones'. The number 'two' originally took the dual, and this usage spread in some Slavic dialects (including the ancestor of Russian) to the following numbers 'three' and 'four' by contamination. (Contamination frequently affects numeric systems: neighboring numerals are notorious for influencing each other phonologically, morphologically, and – as in this case – even syntactically.) The final step in the development was the formal merger of the dual with the genitive singular in some declensions, which led to a nearly complete morphological replacement of the dual by the genitive singular elsewhere in Russian and several of its relatives.

Old Church Slavonic

18.33. Slavic literacy began in 863, when two Greek brothers and missionaries, Constantine and Methodius, arrived in Moravia to teach the Christian faith in Slavic. The brothers had learned the language while growing up in Thessalonika. Constantine devised a script for it, while Methodius did the lion's share of the translation work. Constantine later went to Rome and on his deathbed became a monk, whereupon he took on the name Cyril. The script we call Cyrillic in his honor is not actually the script he invented; the latter is now called Glagolitic (OCS *glagoliŭ* 'word'), with distinctive letters of uncertain origin. What we now call Cyrillic (used nowadays to write Russian, Belarusian, Ukrainian, Bulgarian, Macedonian, Serbian, and various non-Slavic languages of the former Soviet Union) was devised about thirty years later in Bulgaria and based on the Greek majuscule (capital) letters.

18.34. None of the two brothers' translations has survived. The earliest preserved Slavic comes from some of their successors in Bulgaria; this language soon spread as the liturgical language for all the Slavs, and because of its role as the language of Slavic Christendom it is usually called **Old Church Slavonic** or **Old Church Slavic**. OCS is very similar to the Common Slavic that has been reconstructed by linguists; but by all estimates, Common Slavic ceased to exist as a unitary language several hundred years before the time of Constantine and Methodius, and at most OCS can be regarded as one of its later dialects, albeit artificial (that is, it did not quite match the local spoken dialects of the regions in which it was used). The Old Church Slavonic of Bulgaria, regarded as something of a standard, is often called **Old Bulgarian** (or **Old Macedonian**).

Eastern Orthodox Christianity spread to all the East Slavs and the eastern territories of the South Slavs (inhabited by the ancestors of today's Serbs, Macedonians, and Bulgarians), and OCS spread with it; since a different dialect was spoken in each area, OCS was artificially modified to accommodate local dialect features. The term "Church Slavic" is often encountered as a general rubric for these different regional varieties of OCS. Real OCS was not spoken for terribly long before the development of the regional Church Slavic varieties, hence we do not have many documents in it – only about seven major manuscripts in the Glagolitic alphabet, and two in the Cyrillic alphabet from before the twelfth century. The oldest OCS manuscript, from the tenth century, is known as the Kiev Fragments and is written in Glagolitic.

Old Church Slavonic text sample

18.35. The Old Church Slavonic version of Luke 2:4–7, in normalized spelling.

4 vŭzide že Iosifŭ otŭ Galileje, iz grada Nazaretŭska, vŭ Ijudę, vŭ gradŭ Davydovŭ iže naricajetŭ sę Vitleemŭ, zane bęaše otŭ domu i otčistvię Davydova. 5 napisati sę sŭ Mariejo, obręčenojo emu ženajo, sŭštejo neprazdŭnojo. 6 bystŭ že, egda bystetu, isplŭniše sę dŭne roditi ei. 7 i rodi synŭ svoi prŭvenŭi, i povitŭ i, i položi i vŭ ęslaxŭ, zane ne bę ima męsta vŭ obitęli.

4 And Joseph also went up from Galilee, from the city of Nazareth, to Judea, to the city of David, which is called Bethlehem, because he was of the house and lineage of David, 5 to be registered with Mary his betrothed, who was with child. 6 And while they were there the time came for her to be delivered. 7 And she gave birth to her first-born son and wrapped him in swaddling cloths, and laid him in a manger, because there was no place for them in the inn.

tion (§18.20); the pronoun *se* is the old accusative of the reflexive pronoun, from PIE **sye-* (§7.12). *běaše*: 'was', 3rd sing. imperfect; the ending *-aše* is palatalized from *-ax-* before the front vowel *-e* (§18.20). The *-x-* is ultimately from the *s* of the *s*-aorist; this **-s-* changed to *-x-* by the "ruki" rule in certain forms, and then spread to other forms where "ruki" would not have applied, as here. *domu*: 'home, house', with *-u* < **-ou-s-*, the *u*-stem genit. sing. *otičství*: 'lineage', genit. sing.; abstract noun from *otiči* 'father'.

5. *napisati*: 'to register', from *pisati* 'to write'. *sū*: 'with', cp. Lith. *sū*. *obročenojō*: 'betrothed' (instr. sing.), past passive participle of *obročiti* 'to betroth', from *rpka* 'hand' (Russ. *ruka*). *emu*: 'him', dative of the masc.-neut. demonstrative pronoun, containing the oblique case-formant *-m-* (see further §18.69). *ženojō*: 'wife, woman', PIE **g^{en}-* (Gk. *gunē* 'woman', Eng. *queen*). *sōštejō*: 'being', pres. participle < **sont-*; the *-št-* is the South Slavic development of **-tj-*. *neprazdūnojō*: 'pregnant', lit. 'not idle, not empty'.

6. *bystū*: 'it was, it happened (that)', 3rd sing. perfective aorist of *byti* 'to be'. A shorter 3rd singular, *by*, is the historically older form (< **bhuH-s-t*); *bystū* is a blending of *by* and the present 3rd sing. *jestū* 'is' (< **h₂esti*). *byste*: 'both of them were', 3rd dual aorist. *isplūnišē se dīne*: lit. 'the days filled themselves' = 'the days became full, became complete'; *s*-aorist of *isplūniti*, from *plūnū* 'full' (PIE **pl_hno-*, cp. Eng. *full*). *Dīne* is nomin. pl. of *dīnī* 'day', from PIE **dīn-* 'day'. *roditi*: 'to give birth'; *rodī* below is the 3rd sing. aorist. *ei*: 'for her', dat. sing.

7. *synū*: 'son', accus. sing. Usually in animate nouns in Slavic the accusative and genitive singulars are identical, but not here; the genitive is *synu*. The paradigm of 'son' is archaic in this respect, for it comes from a proterokinetic *u*-stem with zero-grade of the suffix in the nominative and accusative (**-u-*, whence Slavic *-ŭ*) and full grade in the other cases (**-eu-* or **-ou-*, whence Slavic *-u*). *svoi*: '(her) own', reflexive adjective, PIE **syo-*, with the **-y-* preserved (but lost in the unstressed pronoun *se* above). *prīvēnīci*: 'first-born', from *prīvū* 'first'. *povītū*: 'wrapped', 3rd sing. aorist; related to Ved. *vāyati* 'weaves', Lat. *uieō* 'I twist, plait'. *i*: 'him', enclitic object pronoun. *položī*: 'lay', PIE **logh-eje-* 'cause to lie', causative of **legh-*. *šlaxū*: 'manger', loc. pl. (sing. in meaning). *bē*: 'was', imperfective aorist of *byti* 'be'. *ima*: 'for them', dat. pl. *mēsta*: 'place', genit. sing.; in Slavic, negative existential arguments require the genitive, so lit. 'there was no(thing) of place'. Compare French *pas de temps* 'no time', literally 'not of time'.

Modern Slavic Languages

East Slavic

18.36. For the first two centuries or so of East Slavic literacy, which began after the early Russian state known as Rus' was Christianized in 988, there were no individually differentiated East Slavic languages, only an East Slavic dialect area – itself only a little different from the rest of Slavic. The capital of Rus' was Kiev, which is where the first East Slavic literature hails from. As part of the Slavic East Orthodox community, the East Slavs adopted Old Church Slavonic as their written liturgical language. It was modified somewhat to incorporate specifically Kievan features and is thus often called Russian Church Slavonic. The literary language has traditionally been called Old Russian, the term we use in this book, but it is a bit of a misnomer: it was ancestral not just to Russian, but also to the other East Slavic languages, Belarusian and Ukrainian. In fact, some features of the "Old Russian" of Kievan Rus' are found in modern Ukrainian but not Russian – not surprising when one recalls that Kiev is now the capital of Ukraine. More accurate terms for Russian Church Slavonic and Old Russian are *Rusian Church Slavonic* and *Old Rusian* (not a typographical error – named after Rus').

The oldest surviving dated East Slavic manuscript is the Ostromir Gospel, dating to 1056–7, but at least one work, the *Sermon on Law and Grace* of Hilarion, was composed slightly earlier in the same century. Principal original compositions of the early Kievan period were hagiographies and historical chronicles. One isolated epic in rhythmic prose, called *Slovo o polku Igoreve* (*The Song of Igor's Campaign*), is known, supposedly composed in the late twelfth century; but its authenticity has been the subject of controversy, and the sole manuscript was destroyed by fire during the War of 1812. If genuine, it would represent the literary highpoint of Kievan Rus'.

18.37. The Tatars conquered Rus' in the 1230s and were not finally overthrown until 1480, after which the northern city of Moscow became the political and literary center of the country. The term Rus' and the adjective "Russian" (Old Russian *Rusŭskyi*) now became identified in the north with the regions around Moscow and Novgorod, the old northern center of trade. As Muscovy grew stronger and gathered the territories that would eventually become known as the Russian Empire, the southern region around Kiev became referred to simply as the "border area," *U-kraina* (from *kraj* 'border'). The Old Russian literary period ended with the rule of Peter I, called the Great (sole ruler 1696–1725), although many date the beginning of Modern Russian to the work of Russia's national poet, Aleksandr Sergeyevich Pushkin (1799–1837), who fixed a number of aspects of the literary language that had been in flux.

18.38. Closely allied with Russian is **Belarusian** (Byelorussian or White Russian), whose standard form is based on the dialect of Minsk. Some old East Slavic texts (traditionally called Old Russian, as per the above) begin to show characteristic Belarusian sound changes in the first half of the thirteenth century. From the fifteenth to the seventeenth century Byelorussia belonged to the Lithuanian and Polish cultural and political world, during which time a wealth of Belarusian literature was produced. It declined as a literary language after that and was not fully resuscitated until after the Russian Revolution.

18.39. Belarusian forms a sort of bridge between Russian and the third East Slavic language, Ukrainian (once called Ruthenian or Little Russian; the latter term is now somewhat offensive). Like Belarusian, Ukrainian was not distinct from Russian until around the thirteenth century. With the fall of Kiev as a center of power, the language of the region practically ceased to be written, although Church Slavonic continued to be cultivated, with increasing admixture of Ukrainian features. A true native Ukrainian literary language did not emerge until the late eighteenth century. Today Ukrainian is spoken by more people than any Slavic language besides Russian.

18.40. All of East Slavic is characterized by the retention of the mobile accent inherited from Common Slavic, but the contrastive pitch has been lost. Most of the East Slavic varieties, with the exception of some northern dialects of Russian and most of Ukrainian, have undergone vowel reduction in unstressed syllables. The strong yers became *e* and *o* (for Russian examples, recall §18.23). The nasal vowels *ę* and *ǫ* were denasalized to *ja* and *u*, as in Russ. *beru* 'I take' (OCS *berę*) and *piat'* 'five' (OCS *pęti*).

West Slavic

18.41. The Slavic language with the third-largest number of speakers is Polish. Texts in this language date to the fourteenth century, although isolated names and

words are found going back to 1136. Since Poland was Christianized by Catholics rather than by East Orthodox missionaries, it has always used the Latin alphabet. The literary works from the fourteenth and fifteenth centuries show strong influence from Czech. To the west and southwest of Gdańsk in northern Poland Kashubian is spoken, viewed by some as a divergent dialect of Polish, by others as a separate language (the issue tends to get politicized). Poland is (or was) also the home of two extinct languages, Slovincian, apparently an old type of Kashubian, and Pomeranian, the latter spoken along the Baltic coast. A third extinct West Slavic language, Polabian, known from a very few writings in the late sixteenth and seventeenth centuries, was spoken along the Elbe River in what is now Germany.

18.42. One Slavic language is still spoken in Germany, in a region called the Lausitz in the southeast not far from Dresden: Sorbian (also called Wendish or Lusatian in older works). Distinguished are two varieties, Upper Sorbian and Lower Sorbian. The combined total of speakers is only about 50,000, few of whom are monolingual. The Sorbs are the last remnant of the large medieval Slavic population that lived between the Elbe and the Oder rivers. A Lower Sorbian translation of the New Testament from 1548 is the first literary example of either language. Sorbian is interesting principally for preserving several archaic features that are otherwise found today only in South Slavic: the dual, the imperfect, and the aorist.

18.43. South of Sorbian territory is found Czech. Like Poland, the Czech-speaking historical region of Bohemia was converted to Catholicism (in the tenth century) rather than East Orthodox Christianity, and the Latin alphabet has been used for writing Czech since its first clear attestation in the fourteenth century. But there were also Slavic missionaries in the tenth century who left behind some Church Slavonic documents with a few Czech dialect characteristics. East of the Czech Republic, in Slovakia, is spoken Slovak, which is quite similar to Czech but much more conservative phonetically; the two languages diverged in the fifteenth century.

South Slavic

18.44. Unlike East and West Slavic, which form a continuous dialect area from Russia westwards into Germany, the South Slavs became separated from the other Slavic peoples starting in the sixth century. The early separation led to some significant differences between South Slavic and the rest of the family. South Slavic has much more faithfully preserved the Common Slavic categories of the imperfect and aorist, which were lost in all of East and West Slavic besides Sorbian (§18.42); and it is the only branch of Slavic having languages that preserve the mobile pitch-accent system.

18.45. After Old Church Slavonic, the first attested South Slavic language is Slovenian (or Slovene), in which are found the first Slavic documents written in the Latin alphabet, from the tenth and eleventh centuries. Nothing more is heard of the language, though, until the end of the fifteenth century. Slovenian preserves the inherited mobile pitch-accent and the dual. In spite of its relatively small number of speakers (c. two million) and the small geographic area over which it is spoken, Slovenian has an unusually large number of quite varied dialects.

18.46. To the southeast of Slovenian is spoken Croatian, which is closely related to Serbian farther to the southeast and to the recently named Bosnian in Bosnia and Hercegovina. Serbian and Croatian are mutually intelligible; but the differences have sometimes been exaggerated for political reasons, and they are written in different

alphabets – Croatian in the Latin alphabet, Serbian in Cyrillic. Because of their mutual intelligibility, Serbian, Croatian, and Bosnian are usually thought of as constituting one language called **Serbo-Croatian**. Serbo-Croatian shares with Slovenian the preservation of the mobile pitch-accent system.

18.47. In the southeast of the former Yugoslavia is Macedonia, home of **Macedonian**. (The same name is given to the language of the ancient Macedonians, which is not connected; see §20.12.) Macedonian was not distinguished from Bulgarian for most of its history. Constantine and Methodius themselves came from Macedonian Thessalonika; their Old Bulgarian is therefore at the same time “Old Macedonian.” No Macedonian literature dates from earlier than the nineteenth century, when a nationalist movement came to the fore and a literary language was established, first written with Greek letters, then in Cyrillic. From the point of view of Bulgaria, Macedonian is simply a west Bulgarian dialect.

18.48. We finally come to Bulgarian itself, whose earliest stage is Old Bulgarian or Old Church Slavonic, as we have already seen. Divergences from the artificial language of Church Slavonic that are due to colloquial developments in Bulgarian begin to appear in written documents in the twelfth and thirteenth centuries.

Bulgarian and Macedonian are sharply divergent from Serbo-Croatian and Slovenian in several respects. The morphology of the noun has undergone striking simplification: except for singular masculine nouns, where a nominative and an oblique case are distinguished, there are no cases in the noun anymore. A definite article has been innovated, appearing as a clitic suffixed to its noun, as in Macedonian *ezik* ‘language’ ~ *ezikot* ‘the language’, *kniga* ‘book’ ~ *knigata* ‘the book’, and *vreme* ‘time’ ~ *vremeto* ‘the time’. In contrast to nouns, the verb is richly inflected in all the inherited Slavic tenses and moods plus two new ones, a future tense and a narrative mood. The narrative mood is used to quote material that the speaker does not know to be factual, often with a sense of mistrust; it has sometimes been called informally a “the hell you say” mood.

Some of the unusual characteristics of Bulgarian and Macedonian are shared with other languages of the Balkans, especially Greek and Albanian. These languages are usually seen as constituting the so-called Balkan speech area, which will be discussed in more detail in the next chapter (§19.5).

Baltic

18.49. In historical times, the Balts have occupied a fairly small patch of land along the Baltic Sea, but their prehistoric extent was much greater. In the late Bronze Age, they stretched from near the western border of Poland all the way east to the Ural Mountains. Baltic river names are found across a large swath of now Slavic-speaking territory in eastern Europe and present-day Russia, as far east as Moscow and as far south as Kiev. To the north the Balts were in lengthy and intimate contact with Finnic tribes, who borrowed hundreds of Baltic words, including agricultural and kinship terms, and words relating to tools and other technologies. From these loans it seems that the Balts had achieved considerable cultural prestige in the region.

Little is known about the prehistoric Balts. In the early to mid-second millennium BC, the inhabitants of the Baltic coast, perhaps but not necessarily ancestors of the Balts, began trading amber with central European tribes in exchange for metals

(which the Baltic region was poor in). Amber, considered magical by many peoples, has been called the gold of the north; the Baltic coast has vast deposits of it from huge pine forests that covered the region 60 million years ago. By around 1600 BC amber was being funneled to Mycenae, Anatolia, and places even more distant deep in Asia. Archaeologists have recovered many non-Baltic Bronze Age artifacts in the Baltic regions that must have come in from the amber trade, including a thirteenth-century-BC statuette of a Hittite deity. Most of the world's best amber still comes from this region.

18.50. The first mention of Baltic peoples in ancient literature is disputed. The Greek historian Herodotus named various tribes in their general vicinity, one of which, the Neuri, he says were chased out of their homeland by an infestation of snakes. This has reminded some of the fact that snake-worship was traditional among the Balts, but the connection is tenuous. A half-millennium later, the Roman historian Tacitus mentions a tribe called the Aisti or Aistii who gathered amber and may well have been Balts; but they could also have been Germanic, since their word for 'amber', *glēsum*, seems most similar to *glass*, a Germanic word. Definitely Baltic are the Soudinoi and Galindai mentioned by Ptolemy the geographer (second century AD); they are known from later medieval records too. In the ninth century we hear first of a tribe called the Bruzi (in Old Bavarian) or Brūs or Burūs (in Arabic, tenth century), the source of the word *Prussia*, the historical region of the western Balts.

18.51. Baltic territory began to shrink shortly before the dawn of the Christian era due to the Gothic migrations into their southwestern territories, and shrank a good deal further as Slavic migrations started around the fifth century AD, continuing into the twelfth century. After 1225, the western Baltic lands (Prussia) became casualties of the aggressive eastward expansion of the Teutonic Order, a German religious and military organization that had been summoned to help Poland fight against the Prussians. Farther to the east, where the Lithuanians lived, the Teutonic Knights were repulsed; the Lithuanians established an empire that rapidly expanded through parts of Russia and Ukraine all the way to the Black Sea. The Grand Duchy of Lithuania lasted from 1362 to 1569 and marked the height of Baltic political might. Only toward the end of this period do the first texts in Baltic languages finally appear.

18.52. It is one of the greatest losses to Indo-European studies that the Baltic languages were written down so late, and only after the native mythologies and traditions had been all but eradicated by the Christian invaders. Luckily, some of them live on in the countless Lithuanian and Latvian folksongs or *dainos* (Lith. *dainos*, Latv. *dainas*) and folktales, many of which preserve Baltic cultural material of great antiquity. Comparatively little of this is known in the West, and these songs have not yet been sufficiently mined for the contributions they could make to reconstructing PIE culture.

18.53. Baltic survives today in the form of two languages, Lithuanian and Latvian, which together comprise the East Baltic subgroup. West Baltic consists of the extinct Old Prussian. We know the names of several other Baltic languages that were spoken in medieval Europe, but no texts have survived in them; however, a fair number of personal names are preserved from Curonian, a medieval East Baltic language that seems to have formed a bridge between Latvian and Old Prussian. Note in passing that Estonian, though geographically "Baltic," is not related and belongs rather to the non-IE Uralic family.

Phonology

CONSONANTS

18.54. Like Slavic, Baltic changed the PIE palatal velars to sibilants. Outside Lithuanian, the treatment is the same as Slavic, with *s* and *z* the outcome of PIE **k̑* and **ĝ(h)*: Latv. *sirds*, OPruss. *seyr* 'heart' < **k̑rd-*; Latv. *zirnis*, OPruss. *syrne* 'grain' (with *s-* probably representing *z* in the German-based orthography of the language) < **ĝrh₂-n-* (cp. Lat. *grānum*, Eng. *corn*). In Lithuanian, however, the outcomes are *š* and *ž*: *širdis* 'heart' and *žirnis* 'pea'.

18.55. The "ruki" rule changed an old **s* to *š* in some words in Lithuanian: *viršūs* 'high' (cp. OCS *vȓxŭ* 'high' < **u̯rs-*), *aušrà* 'dawn' < **h₂eus-r-* (cp. **h₂eus-ōs-* in Lat. *aurōra* 'dawn'). As there are many words where the change apparently did not happen, such as *klausyti* 'to hear' and *visas* 'all', and since the change did not affect Latvian or Old Prussian, the change appears not to have spread through the whole of Baltic territory.

SHORT VOWELS

18.56. Short **i* and **u* remained intact: Lith. *likti* 'to leave', Latv. *likt* < **likʷ-* (cp. Lat. *re-lic-tus* 'left behind'); Lith. *šūnis*, Latv. *suns*, OPruss. *sunis*, all genitive singular meaning 'of (a) dog' (PIE **k̑un-os*, cp. Ved. *śúnas*, Gk. *kunós*). Short **e* generally stayed intact as well, as in the words for 'honey', Lith. *medūs*, Latv. *medus*, and OPruss. *meddo* < **medhu-*. But there are a number of problematic words in which **e* is continued by *a*, especially in word-initial position: **eĝ-* 'I' > Lith. *àš*, and **ek̑uo-* 'horse' > OLith. *ašva* 'mare' are two examples. These forms vary dialectally.

18.57. As in Germanic, short **a* and **o* merged as *a*: Lith. *akis*, Latv. *acs*, OPruss. *ackis* (nomin. pl.) 'eye' < **okʷ-* (earlier **h₂ekʷ-*); Lith. *akis* 'axis', Latv. *ass*, OPruss. *assis* < **aĝ-s-* (earlier **h₂eĝ-s-*). Note this merger is opposite that of Slavic (§18.22).

LONG VOWELS

18.58. Long **ā* and **ō* remained distinct in Baltic, in contrast to Slavic. The long **ā* remained intact in Latvian (as in *māte* 'mother' < **māter-*), but became *ō* in Lithuanian (*mótė*). Long **ō* has different reflexes in the different branches of Baltic. It became the diphthong *uo* in East Baltic (Lith. *dúoti* 'to give', Latv. *duot* 'to give' < **dōti* < **deh₃-tei*), but not in West Baltic (contrast OPruss. *datwei* 'to give'). The other long vowels remained intact: Lith. *bėgti* 'to run', Latv. *bēgt* < **bhēgʷ-tei*; Lith. *gývas* 'alive', Latv. *dzīvs*, OPruss. masc. accus. pl. *gi(j)wans* or *geywans* < **gʷī-uo-* (earlier **gʷih₃-uo-*); Lith. *būti* 'to be', Latv. *būt*, OPruss. *buton* < **bhu-* (earlier **bhuH-*).

Morphology

VERBS

18.59. While Slavic preserved the aorist but lost the *s*-future, Baltic fully preserved the *s*-future but lost the aorist. Like Slavic and several other branches of IE, Baltic

created a new imperfect. In the present tense of thematic verbs, the *o*-grade of the thematic vowel has been generalized in all persons: Lith. 3rd sing. *sùka*, 2nd pl. *sùkate* 'turn' from **-o* **-o-te* (cp. the Tocharian verb classes III and IV, §17.28; contrast Gk. *phérei phérete*). A curious and not fully explained trait of Baltic is that in all tenses of the verb, the 3rd dual and plural endings are the same as the 3rd singular: thus Lith. *èsti* means 'is', 'both of them are', and 'they are'.

NOUNS

18.60. The Baltic noun is remarkably conservative, preserving the copious case system inherited from PIE. (See the Lithuanian paradigms given below.) Nouns with stems in *-ê-*, like Lith. *môtê* 'mother', became very prolific in Baltic; their origin is the subject of debate since PIE did not have such a class. Baltic also saw a profusion of new *i*-stems; these developed from old consonant stems by generalizing the vowel *-i-* in the accusative singular **-im*, the outcome of the old consonant-stem ending **-m*.

Lithuanian

18.61. Lithuanian literature begins around 1525, when translations of the Lord's Prayer, the Creed, and the Ave Maria were composed. The language of this period is called **Old Lithuanian**, and differs from the modern language chiefly in inflection: three noun cases were in use that have since become obsolete, the allative, illative, and adessive; and athematic verb inflection was considerably more widespread than it is today. (The extra noun cases were an invention of East Baltic; the illative is also found in older Latvian. They probably arose under the influence of neighboring Finnic languages.) The modern literary language is based on that of the author Jonas Jablonskis (1861–1930), who wrote in a western dialect. Some non-standard dialects preserve features that were in Old Lithuanian but have been lost in the standard language, such as the additional cases in the noun.

18.62. Lithuanian does not differ significantly from the picture drawn above for Common Baltic. The accent is mobile, with two types of intonation (rising and falling) distinguished in long syllables. A long syllable contains a long vowel, a diphthong, or a sequence of vowel plus resonant. The rising intonation is denoted by a circumflex accent (*˘*): *mēdis* 'tree', *sūnai* 'son!', *šūñs* 'of the dog'. The falling intonation is denoted by either an acute (*˙*) or a grave (*˘*) accent: *výras* 'man', *pakl̃austi* 'to ask', *pilnas* 'full'. The grave accent is also used over short vowels to indicate stress but without any particular pitch contrast, as in *iki* 'until'. (Confusingly, the terminology and the accent-marks of Lithuanian are exactly backwards from the rest of Balto-Slavic since apparently Lithuanian flipped the original system: the Lithuanian "acute," though historically from the Balto-Slavic acute, is a falling intonation rather than a rising one, and the Lithuanian "circumflex," historically from the Balto-Slavic circumflex, is rising; compare §18.11.)

18.63. The paradigms below illustrate the accentual classes of Lithuanian nouns. The nouns are *výras* 'man' (PIE *o*-stem; first accent class), *mēdis* 'tree' (PIE **-iō*-stem; second accent class); *rankà* 'hand, arm' (*ā*-stem; second accent class); *sūnūs* 'son' (*u*-stem; third accent class); *šuō* 'dog' (*n*-stem; fourth accent class).

Singular

N	<i>výras</i>	<i>mēdis</i>	<i>rankà</i>	<i>sūnūs</i>	<i>šuō</i>
G	<i>výro</i>	<i>mēdžio</i>	<i>raňkos</i>	<i>sūnaūs</i>	<i>šunš</i>
D	<i>výrui</i>	<i>mēdžiui</i>	<i>raňkai</i>	<i>sūnui</i>	<i>šunuiui</i>
A	<i>výra</i>	<i>mēdj</i>	<i>raňka</i>	<i>sūnų</i>	<i>šunį</i>
I	<i>výru</i>	<i>mēdžiū</i>	<i>rankà</i>	<i>sūnumi</i>	<i>šunimi, šuniū</i>
L	<i>výre</i>	<i>mēdyje</i>	<i>raňkoje</i>	<i>sūnujė</i>	<i>šunyjė</i>
V	<i>výre</i>	<i>mēdi</i>	<i>raňka</i>	<i>sūnaū</i>	<i>šuniē</i>

Plural

N	<i>výrai</i>	<i>mēdžiai</i>	<i>raňkos</i>	<i>sūnus</i>	<i>šunys, šunes</i>
G	<i>výrų</i>	<i>mēdžių</i>	<i>raňkų</i>	<i>sūnų</i>	<i>šunų</i>
D	<i>výrams</i>	<i>mēdžiams</i>	<i>raňkoms</i>	<i>sūnūms</i>	<i>šunims</i>
A	<i>výrus</i>	<i>mēdžius</i>	<i>rankàs</i>	<i>sūnus</i>	<i>šunis</i>
I	<i>výrais</i>	<i>mēdžiais</i>	<i>raňkomis</i>	<i>sūnumis</i>	<i>šunimis</i>
L	<i>výruose</i>	<i>mēdžiuose</i>	<i>raňkose</i>	<i>sūnuosė</i>	<i>šunysė</i>
V	<i>výrai</i>	<i>mēdžiai</i>	<i>raňkos</i>	<i>sūnūs</i>	<i>šunys</i>

The conservativeness of the Baltic noun is easily illustrated by the paradigms above. Thus the singular declension of 'man' faithfully continues all the IE *o*-stem singular case-endings except the genitive singular. (The PIE genitive was lost, and the Baltic ending is from the ablative.) The hooked vowels in forms like *výra* and *šunį* were once nasal vowels that continued sequences of vowel plus nasal.

18.64. It is popularly said that Lithuanian is the oldest or most conservative Indo-European language now spoken. This impression rests largely on the high degree of faithfulness with which it has preserved the aspects of PIE phonology and nominal morphology discussed above, and is bolstered by such close equivalencies between it and Sanskrit as Lith. *kàs* 'who' and Ved. *kás* 'who'. While calling it the "oldest" IE language is a misnomer, its conservativeness in these areas cannot be gainsaid, and probably does exceed that of all the other contemporary IE languages (although such things are not easily quantified). It should be remembered, however, that the language has not been equally conservative in all domains, especially in the verb.

Lithuanian text sample

18.65. An example of a *raudà* or traditional song of lament, in this case on the occasion of the loss of a son. From the collection by A. Juškevič, *Lietuvizkos dainos. užrašytos* (Kazan, 1881-3) III 1192. The spelling has been modernized. Virtually every word in the first stanza goes back to PIE.

Ō māno sūnėli, māno dobitėli, ō māno artojęli, māno šienpiūvėli! Išauš pavasarėlis, visų sūnėliai pō lėgius laukeliūs sū naujomis žagrėlėmis, sū šemaīs jautėliais; ō žmonių sūnėliai pō lėgius lankelės sū šviesais dalgėliais švytūja.

Ō māno sūnėlio būbūja šėmi jautėliai, sāvō artojėlio pasigėnda; ō māno vaikėlio rūdýja šviēsūs dalgėliai.

Išauš pavasarėlis; žmonių moterėlis pō ulytėlės sū glėbūciais nešiójas; ō aš vienà, ō aš vienà verkiū pasižiūrėdama; ō aš verkiū, neturiū nei jokiōs patiekėlės, neturiū sū kūmi pasidžiaugti.

Ō māno sunēli, māno diemedēli, kād āš būčiau užsiauginus, aīt kojēliņ pastāčius.
Ō māno mažiukēli, aī māno rankēles taiņ suīkios, aī māno žodēliai taiņ skaūdūs?

O my little son, my little clover, o my little plowman, my little hay-mower! The springtime will grow light, everyone's sons (will be) in the flat fields with new plows, with young gray oxen; o men's dear sons make shiny motions with their gleaming scythes in the flat meadows.

O the young gray oxen of my little son bellow, they long for their little plowman; o the gleaming scythes of my little boy grow rusty.

Spring will grow light; the mother of men (will go about) in the streets carrying (her children) with her embraces; o I (am) alone, o I alone cry while looking about; o I cry, I may not take consolation in anything, may not delight in anyone.

O my little son, my little southernwood, if only I had reared (you), had placed (you) on (your) little feet. O my little one, (was) my hand too heavy, (were) my words too painful?

18.65a. Notes (to first stanza only). *māno*: 'my', dialectal for standard *māno*. *sūnēli*: 'little son', vocative of *sūnēlis*, diminutive of *sūnūs*, cognate with Eng. *son* (PIE **suHnu-*). Practically every noun is turned into a diminutive in this lament, with the suffix *-ēlis* (or *-ēlis* when the base has two syllables); diminutives are frequently used to express tenderness or other emotive shades of meaning. They have sometimes been rendered as 'little' or 'dear' in the translation above, and sometimes not translated in any special way. *dobilēli*: 'little clover', vocative diminutive of *dōbilas*. *artojēli*: 'little plowman', vocative diminutive of *artōjis*, agent noun from *arti* 'to plow' < PIE **h₂erh₃-*, cp. Lat. *arāre* 'to plow'. *šienpiūvēli*: 'little hay-maker', dialectal voc. diminutive of *šienpiūvis*, from *šiēnas* 'hay' (cp. OCS *sēno*) and *piāuti* 'to cut, mow' (cp. Lat. *pū-tāre* 'to cut'). *išauš*: 'will become daylight, will become bright day', 3rd sing. future of *iš-aušti*; cp. *aušrā* 'dawn' < PIE **h₂eus-*. *pavasariēlis*: 'little springtime', from *pavāsaris* 'spring', lit. 'upon-summer' (*vāsara* 'summer'); cp. Czech *podzim* 'fall', lit. 'upon-winter' (*zima* 'winter'). *visū*: 'of all', genit. pl. of *visas*; cognate with OCS *visi* 'all, entire', Ved. *vis-va-* 'every'. *laukelius*: 'little fields', accus. pl. diminutive of *laukas*, cognate with Lat. *lūcus* 'grove' and Eng. *lea* 'meadow'. *sū*: 'with', cp. OCS *sū*. *naujomis*: 'new', instr. pl. of *naūjas*, from PIE **neu-jo-* (also in Ved. *nārya-*), a by-form of **neyo-* (Hitt. *nēwa-*, Lat. *nōnus*, and Eng. *new*). *šēmaīs*: 'blue-gray' (said of animals), instr. pl., related to Ved. *šyāmā-* 'dark'. *jautēliais*: 'young oxen', instr. pl. diminutive of *jautis* 'ox', cp. Ved. *yuvāti* 'binds'. *žmoniņ*: 'of men', genit. pl. The paradigm is irregular; the singular is *žmogūs*, the plural is *žmōnēs*, which is morphologically feminine and comes from the stem of the Old Lithuanian *n*-stem singular *žmuo*. The word is cognate with OPruss. *smunu(n)* 'person', Lat. *homō* 'man, person' (stem *homin-*), from PIE **dhǵh(ē)mō(n)* (cp. §6.30). *šviesais*: 'shining, gleaming', instr. pl.; from PIE **kuei-s-*, related to *švytūja* at the end of the sentence. *švytūja*: 'flash, make a gleaming motion', dialectal 3rd pl. of *švytioti*, from **kūit-*, a neo-lengthened grade (§18.13) from **kūit-*, zero-grade of **kuei-t-* (the root of OCS *svītiti* *sg* 'to shine' and Eng. *white*).

Latvian

18.66. Latvian, like Lithuanian, is first attested in a translation of a catechism, dating to 1585. The two languages are quite similar, but Latvian is less conservative and gives the impression of a Lithuanian that has been somewhat slimmed down. The accent has been fixed on the initial syllable, which has led to widespread apocope of vowels in final syllables: thus contrast Latv. *rīts* 'day' and *būt* 'to be' with Lith.

rýtas and *búti*. Unlike Lithuanian but like Old Prussian and Slavic, the PIE palatal stops **k̑* and **ǵ(h)* became *s* and *z*: *desmit* ‘ten’ (**dek̑m-t-*; cp. Lith. *dėšimt*), *ziema* ‘winter’ (**ǵheim-*; cp. Lith. *žiema*). Latvian has lost *n* before consonants, as in *ruoka* ‘hand’ versus Lith. *rankà*. The consonants have been widely affected by palatalization, as in *acis* ‘eyes’ (Lith. *akis*) and *dzērve* ‘crane’ (Lith. *gervė*). In the native orthography, palatalization is often indicated with a comma underneath the letter (as in *ceļš* ‘road’ in the text passage below).

In morphology, the language is similar to Lithuanian; its most characteristic innovation is a mood known as the debitive, expressing necessity and formed by prefixing the syllable *jā-* (from the relative pronominal stem **ǵo-*) to a 3rd singular verb form: *man jālasa grāmata* ‘I must read a book’. The construction is impersonal; the logical subject (‘I’, *man*) is in the dative case, and the logical object (‘book’, *grāmata*) is in the nominative.

Latvian text sample

18.67. Excerpt from the poem “Atraitnes dēls” (“The Widow’s Son”) by Vilis Plūdons (1874–1940).

- Es kāpt gribu kalnā visaugstākā,
 Ne spēka man pietrūks, ne bail man no kā,
 Lai putenis plosās, lai ziemeļi elš,
 Uz augšu, uz augšu tik ies man ceļš.
- 5 Es laidīšos jūras dziļumos,
 Un meklēšu, pētīšu, vandišu tos,
 Līdz pērli būšu tur atradis,
 Par visām skaistāk kas mirgo un viz.
- I want to climb the highest mountain,
 I shall neither lack any strength, nor shall I fear anything,
 Although the blizzard rages, although the north wind howls,
 Still my road will climb up and up.
- 5 I shall dive into the depths of the ocean,
 And I shall look for, explore, (and) rummage through them
 Until I shall have found a pearl there,
 Which glitters and gleams more beautifully than anything.

18.67a. Notes. 1–2. *es*: ‘I’; in contrast to Lith. *aš* it preserves the inherited **e* of PIE **eǵ-*. *kāpt*: ‘to climb’; *-t* is the infinitive ending (Lith. *-ti*, Baltic **-tei*). *gribu*: ‘I want’, 1st sing. present of *gribēt* ‘to want’. *kalnā*: ‘mountain’, locative sing. *visaugstākā*: ‘highest’, superlative of *augsts* ‘high’. *ne ... ne*: ‘neither ... nor’. *spēka*: ‘strength’, genit. sing. of *spēks* ‘strength’, object of the impersonal verb *pietrūks*, which takes a genitive object and a dative subject (*man*, ‘I’; lit. ‘to me there will not be wanting of strength’). *pietrūks*: ‘will lack’, 3rd sing. future; *-s* is the future suffix. *bail*: ‘fear’. *no kā*: ‘from anything’; *kā* is genitive of *ka* ‘something, anything’.

3–6. *lai*: ‘although’. *putenis*: ‘blizzard’. *plosās*: ‘rages’, 3rd sing. of the reflexive verb *plosīties*; in both forms the *-s* is the reflexive particle, from PIE **syē-*. *ziemeļi*: ‘north wind’, from *ziema* ‘winter’. *elš*: ‘howls’, 3rd sing. of *elsāt*. *uz*: ‘up’. *augšu*: ‘I will climb’, 1st sing. future of *augt*. The *-š* is palatalized from *-s*, as also in the forms *ceļš*, *laidīšos*, *meklēšu*, *pētīšu*, *vandišu*,

bušu below. ies: 'will go', 3rd sing. fut. celš: 'road'. laidšos: 'I will dive into', reflexive verb. jūras: 'ocean', genit. sing. of *jūra*. dziļumos: 'depths', accus. pl. of *dziļums*, from *dziļs* 'deep'. meklēšu, pētīšu, vandīšu: all 1st sing. future. tos: accus. pl. demonstrative pronoun.

7–8. līdz: 'until'. būšu tur atradis: 'I shall have found', periphrastic future perfect of *atrast* 'to find'. par: 'than', preposition, PIE **per* 'through'. visām: 'everything', cp. note on Lith. *visū* above (§18.65a). skaistāk: 'more beautiful', comparative of *skaists*. Note that this whole phrase has been fronted before the relative pronoun *kas*, a syntactic feature characteristic of poetic language. *kas*: 'which'. mirgo: 'glitters', 3rd sing. present of *mirdzēt*. viz: 'gleams', 3rd sing. of *vizēt*.

Old Prussian

18.68. Old Prussian is the only known West Baltic language, and unfortunately it is not known all that well. It was spoken in the former German territory of East Prussia (now a small area of northeastern Poland along the Gulf of Gdańsk and extending to Kaliningrad), which became German when it was conquered by the knights of the Teutonic Order in 1283. We possess two vocabularies and three texts in Old Prussian. First is the Elbing Vocabulary, compiled around 1300 but existing in a copy from a century later; it contains 802 words from a southern dialect. Of poorer quality is a list of 100 Prussian words in the Prussian Chronicle of a monk named Simon Grunau from the early 1500s; these are based on northwestern Old Prussian dialects. Our only connected texts are three Lutheran catechisms from the mid-1500s that are translated from German. The translations are not of high quality, and there is much inconsistency in orthography; but they do provide us with our sole window onto the grammar and structure of the language. We also possess many Prussian personal names from German sources. By the close of the seventeenth century the language had become extinct.

18.69. Old Prussian is of great interest because it shares several traits with Slavic but not Baltic, or with other IE languages but not the rest of Balto-Slavic. For example, while the rest of Balto-Slavic replaced the *o*-stem genitive singular with the ablative, Old Prussian has a form in *-as* which can only come from PIE **-os(i)o* like Skt. *-asya* and Homeric Gk. *-oio*. While the rest of Balto-Slavic has a present passive participle in **-mo-*, OPruss. preserves the more familiar type in **-m(e)no-* or **-mh₂no-* (§5.56), as in the form *poklausīmanas* '(are) heard'. The pronominal oblique stem-formant **-sm-* that is found in Indo-Iranian and Italic (§7.8) also makes an appearance in Old Prussian forms like the dative *stesmu* 'to the' (cp. Ved. dat. *tásmai*, South Picene locative *esmen*), while the rest of Balto-Slavic has only *-m-* (e.g. Russ. *tomu* 'to him', Lith. *tāmui* 'to him', Latv. *tam*). Old Prussian is also the only Balto-Slavic language to preserve the original *n-* of the word for 'nine' (spelled *newints* or *newyntz*); in the rest of Balto-Slavic, the **n-* was replaced prehistorically with a *d-* that spread from the following number 'ten' (cp. OCS *devetī*, Lith. *devyni*, alongside OCS *desetī* 'ten', Lith. *dešimt*). One wonders what other archaic features the language boasted of which we have no knowledge.

Old Prussian text sample

18.70. The Ten Commandments from the Second Catechism of 1548.

Staeý dessimpts pallapsaey

Pirmois. Tou ni tur kittans deiwans turryetwey.

Anters. Tou ni tur sten emnen twayse deýwas ni enbaenden westwey.

Tirtis. Tou tur stan lankinan deýnan swyntintwey.

Ketwirtz. Tou tur twayien thawán bhae mutien smunintwey.

Pyenkts. Tou ni tur gallintwey.

Usts. Tou ni tur salobisquan limtwey.

Septmas. Tou ni tur ranktwey.

Asmus. Tou ni tur reddi weydikausan waytaton preyken twayien tauwyschen.

Newyntz. Tou ni tur pallapsitwey twaysis tauwyschis butten.

Dessympts. Tou ni tur pallapsitwey twaysis tauwyschies gennan, waykan, mergwan, pecku, adder ka-tanaessen hest.

The Ten Commandments

First. Thou shalt not have other gods. Second. Thou shalt not take the name of thy God in vain. Third. Thou shalt keep holy the sabbath. Fourth. Thou shalt honor thy father and mother. Fifth. Thou shalt not kill. Sixth. Thou shalt not commit adultery. Seventh. Thou shalt not steal. Eighth. Thou shalt not bear false witness against thy neighbor. Ninth. Thou shalt not covet thy neighbor's house. Tenth. Thou shalt not covet thy neighbor's wife, manservant, maidservant, cattle or what is his.

18.70a. Notes (selective). *staeý*: 'the', definite article, nomin. pl.; compare the accusative singulars below, *sten* and *stan*. *dessimpts*: 'ten', cp. Lith. *dėšimt*. *pallapsaey*: 'commandments', cp. Lith. *liėpti* 'to order'. *pirmois*: 'first', cp. Lith. *pirmas*; see §7.18. *tou*: 'you', cp. Lith. *tù*; lengthened **tū* in OCS *ty*. *ni*: 'not'. *tur*: 'have to, must', cp. Lith. *turėti* 'have to'. *kittans*: 'other', accus. pl.; cp. Lith. *kitas* 'other'. *deiwans*: 'gods', accus. pl., exactly continuing PIE **deiy-ons* (cf. §6.55). *turryetwey*: 'have', infinitive; same verb as *tur* earlier in the sentence. The infinitive ending *-twey* is similar to the infinitive ending *-ti* (< **-tei*) in the rest of Balto-Slavic, which goes back to the locative of a verbal abstract noun in **-ti-*. The OPruss. form appears to go back to an abstract in **-tu-* instead (§5.54).

anters: 'second', literally 'the other', and cognate with Eng. *other*. *emnen*: 'name', metathesized from **enmen*; Slavic has **ime* from earlier **inmen-* from **h₁n(h₂)-men*; see §6.36. *westwey*: literally 'to lead', from PIE **ueǵh-* 'convey' (> OCS *vezti* 'to convey', Lat. *uehō* 'I carry, convey', also in Eng. *wagon*). *tirtis*: 'third'. *deýnan*: 'day', accus. sing. *swyntintwey*: 'keep holy', cp. Lith. *šventas* 'holy', OCS *svetŭ* 'holy', all cognate with Avestan *spənta-* 'holy'; PIE **k_uento-*. *ketwirtz*: 'fourth'. *thawan*: 'father', accus. sing.; cp. Lith. *tėvas*. *mutien*: 'mother', accus. sing., cp. Lith. *mótina*.

pyenkts: 'fifth'. *gallintwey*: 'to kill'. *usts*: 'sixth', written *wuschts* in the First Catechism. This would appear to come from **uk-to-*, with **uk-* being some kind of reduced form of PIE **sueks* 'six'. Its closest relative is dialectal Lith. *ušės* 'puerperium' (the period of six weeks of rest after giving birth). *salobisquan*: 'marital'; ultimately a borrowing from Slavic (cp. Polish *ślubić* 'to get engaged'). *limtwey*: 'to break'; cp. OCS *lomiti*. *septmas*: 'seventh', closer to PIE **sept(m)mo-* and OCS *sedmŭ* than Lith. *septiūtas* and Latv. *septiņš*. *ranktwey*: 'to steal'. *asmus*: 'eighth'; cp. Lith. *āšmas*, OCS *osmŭ* < **okt-mo-*, probably related from whatever the PIE form was (see §7.20) under the influence of the preceding **sept(m)mo-*. *reddi*: 'false'. *weydikausan*: 'witness'. *tauwyschen*: 'neighbor'; cp. Latv. *tuvs* 'near'.

newyntz: 'ninth', important in preserving the *n-* (see §18.69). *pallapsitwey*: 'covet'. *butten*: 'house'; probably from the same root as German *bauen* 'to build'. *dessympts*: 'tenth'. *gennan*: 'wife', PIE **g^wen-* 'woman' (> Eng. *queen*). *waykan*: 'manservant', accus. sing. of *waix*. *pecku*: 'cattle', virtually unchanged from PIE **peku*, but problematic in showing a centum-like

treatment of the palatal *k*. See §18.5. *ka-*: 'which', relative/interrogative pronominal stem (PIE **k^wo-*), cliticized to the following word. *hest*: 'is'; the *h-* is unexplained. It is usually spelled *est* or *aest* (*ast* in the First Catechism).

For Further Reading

There is no general reference book treating historical Balto-Slavic grammar as a unit; most books treat either Baltic or Slavic, or treat just one aspect of Balto-Slavic. In the latter category may be chiefly mentioned Stang 1942, which treats the verb. Two highly technical works, Kurylowicz 1958 and 1956 (see Bibliography, chs 3 and 4), are on general topics in IE phonology but have a wealth of material on the Slavic and Baltic vowel systems and accentology.

The most thorough comparative grammar of Slavic is Vaillant 1950–77, which also treats the developments in the Slavic daughter languages. Shorter and also very useful is Meillet 1965 (with revisions and expansions by Vaillant). The history of Slavic phonology from PIE to the modern languages is given detailed and able treatment in Carlton 1990, although the few pages on PIE itself are a bit old-fashioned. On Slavic accentology, see also Stang 1965. No comprehensive Slavic etymological dictionary is yet complete; Trubachev 1974–, arranged by reconstructed Slavic form, is now up to *obsojŋisa*. The standard etymological dictionary of Russian is Vasmer 1953–8; a revised edition appeared a few years later in Russian, and still has to do double duty as a Slavic etymological dictionary. The best grammar of Old Church Slavonic in English is Lunt 2001, although Indo-Europeanists tend to favor Leskien 1990.

The history of Baltic phonology and morphology is masterfully covered in Stang 1966, but its accessibility is lessened by the fact that quotes from scholarly literature written in various Slavic and Baltic languages are left untranslated. A supplementary volume with an index and corrections was published in 1975. Easier to use (and in English) is Endzelīns 1971. Lithuanian is well served with a good etymological dictionary, Fraenkel 1962–5; Zinkevičius 1996 has some good sections on historical grammar. Note also Senn 1957–66, a classic reference grammar whose second volume has sample texts and a glossary. For Old Prussian, useful and accessible is Schmalstieg's 1974 historical grammar; the texts are collected in Trautmann 1910. Two etymological dictionaries of Old Prussian are (in Lithuanian) Maziulis 1988–97 and (in Russian) the incomplete Toporov 1975– (through L).

For Review

Know the meaning or significance of the following:

"ruki" rule	Law of Open	Third Palatalization	Constantine and
pitch-accent	Syllables	yer	Methodius
Winter's Law	First Palatalization	ToFT groups	Rus' <i>dainos</i>
"neo-lengthened"	Second	perfective	
grade	Palatalization	imperfective	

Exercises

- 1 Explain the history or significance of each of the following Slavic forms. The forms are in OCS unless otherwise noted.

a	<i>tysęšćī</i>	e	<i>uxo</i>	i	Russ. <i>gorod</i>
b	<i>sridice</i>	f	<i>usūnptī</i>	j	<i>peřī</i>
c	Russ. <i>umirat'</i>	g	<i>nosū</i>	k	Old Russ. <i>byšęšćī</i>
d	<i>bogū</i>	h	<i>myšī</i>	l	Maced. <i>knigata</i>

- 2 Explain the history or significance of each of the following Baltic forms. The forms are Lithuanian unless marked otherwise.

a	<i>sęsti</i>	e	<i>vīķas</i>	i	Latv. <i>ziema</i>
b	<i>klausyti</i>	f	<i>viršūs</i>	j	OPruss. <i>poklausimanas</i>
c	<i>širftas</i>	g	<i>akis</i>	k	OPruss. <i>newints</i>
d	<i>pilnas</i>	h	<i>dūoti</i>		

- 3 Give the PIE velar from which the boldfaced segments in the following forms descend. In the Slavic forms, there has been no palatalization. Some forms may have more than one answer.

a	Serbo-Croatian <i>zūb</i> 'tooth'	f	Latv. <i>šīds</i> 'heart'
b	Lith. <i>káršti</i> 'to age'	g	Lith. <i>gérvė</i> 'crane'
c	Pol. <i>krew</i> 'blood'	h	Lith. <i>kàs</i> 'who'
d	OCS <i>slava</i> 'fame, glory'	i	Latv. <i>zīrnis</i> 'grain'
e	Lith. <i>žmogūs</i> 'person'		

- 4 For each of the following reconstructed proto-Slavic forms, identify whether it would later be affected by one of the Slavic palatalizations, and if so, by which one.

a	* <i>rpkę</i> 'hands' (nom. pl.)	d	* <i>sněgīnū</i> 'snowy'
b	* <i>nogę</i> 'legs' (nom. pl.)	e	* <i>starīkī</i> 'old man' (nom. sing.)
c	* <i>brěgī</i> 'river bank' (loc. sing.)		

- 5 Identify the yers in each of the following OCS words, and for each yer indicate whether it is weak or strong.

a	<i>vīsī</i> 'village'	e	<i>vūzměriti</i> 'to measure'
b	<i>dvīrīnikū</i> 'doorkeeper'	f	<i>sūnīnū</i> 'relating to sleep'
c	<i>dlīgotrīpělīvū</i> 'long-suffering'	g	<i>otūdesęětīvovati</i> 'pay a tithe'
d	<i>zūvati</i> 'to call'	h	<i>vūgodīnikū</i> 'servant'

- 6 Explain the role that Balto-Slavic plays in the laryngeal theory.
- 7 In very basic terms, how do accent patterns in a Lithuanian consonant-stem noun paradigm like the following differ from accent patterns in PIE consonant-stem paradigms?

nom. sg.	dukt ė 'daughter'	sesu o 'sister'
dat. sg.	dūkt er iai	sęser iai

- 8 A sound change known as Saussure's Law affected the accentuation of certain words in Lithuanian. Observe the forms below of the nouns for 'floor', 'zero', and

'man'. The last column contains forms of the definite adjective meaning 'the white ...'.

instr. sg.	<i>aukštù</i>	<i>nuliù</i>	<i>výru</i>	<i>baltúo-ju</i>
accus. pl.	<i>aukštùs</i>	<i>nuliūs</i>	<i>výrus</i>	<i>baltúos-ius</i>
genit. pl.	<i>aūkštų</i>	<i>nūlių</i>	<i>výrų</i>	<i>baltų-jų</i>
instr. pl.	<i>aūkštais</i>	<i>nūliais</i>	<i>výrais</i>	<i>baltaĩs-iais</i>

The intonations on the endings of the definite adjective (before the hyphens, which are added for clarity) can be thought of as the underlying intonations for these endings in all the forms above. ´ is an acute on a long vowel or diphthong, ` is an acute on a short.

- In the nouns, come up with a rule for when a stem is accented and when an ending is accented.
- The rule you came up with in **a** is due to a sound change. Assuming that the accent was originally on the stem in all forms of the nouns, state the sound change.
- Does this sound change also account for the accentuation of the adjective forms? Explain.
- Both the *-uo-* and the *-u* of *baltúoju* and *baltúosius* (and the *-u* of the corresponding noun forms) come from **ō*. State a sound change to account for the distribution of these two outcomes of **ō*.
- If the verb *māto* 'sees' has infinitive *matyti*, what would the infinitive of *ródo* 'shows' be?

PIE Vocabulary X: Time

- *nokʷt-* 'NIGHT': Hitt. *nekuz* 'of eventide', Gk. *núks* (*núkt-*), Lat. *nox* (*noct-*), Lith. *naktis*
**h₂eusōs* 'dawn': Ved. *uṣás*, Gk. (Doric) *āōs*, Lat. *aurōra*
**ǵheim-*, **ǵheim-* 'winter': Gk. *kheĩma*, Lat. *hiems*, OCS *zima*
**yes-ǵ* 'spring': Gk. *éar*, Lat. *uēr*, OCS *vesna*, Lith. *vāsara*
**uet-* 'year': Hitt. *witt-*, Gk. (*w*)*étos*, Lat. *uetus* 'old'
**jeh₁-ro-* 'YEAR': Av. *yārə*, Gk. *hōrā* 'season'

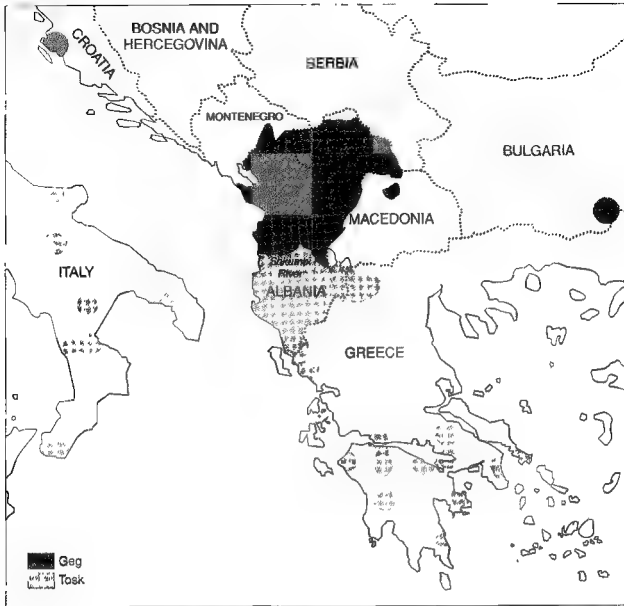
19 Albanian

Introduction

19.1. Albanian is the official language of Albania, where it is spoken by over 3 million people. Almost as many speak it in various areas of the former Yugoslavia, and smaller but not insignificant numbers of speakers live in enclaves in Macedonia, Greece, and southern Italy (Calabria and Sicily). The speakers in northern Albania, the former Yugoslavia, Macedonia, and Turkey speak the dialect called **Geg** (or **Gheg**, Albanian *Gegë*), while Albanians living in southern Albania, Greece, and Italy speak **Tosk** (*Toskë*), of which the Italian variety, called *Arbëresh* (Albanian, etymologically), is the most archaic. In Albania, the two dialects are separated roughly by the Shkumbin River. Both dialects have a literary history (see below); the official variety of Albanian since 1952 is Tosk, but with a goodly amount of admixture from Geg. The two dialects are mutually intelligible in their standard varieties, although numerous subdialects exist that show considerable variation, especially in the north and northeast of the Geg-speaking area. The Albanian-speaking populations of Greece and Italy arose probably between the thirteenth and fifteenth centuries.

19.2. Albanian forms its own separate branch of Indo-European; it is the last branch to appear in written records. This is one of the reasons why its origins are shrouded in mystery and controversy. The widespread assertion that it is the modern-day descendant of Illyrian, spoken in much the same region during classical times (see §§20.13ff.), makes geographic and historical sense but is linguistically untestable since we know so little about Illyrian. Competing hypotheses, likewise untestable, would derive Albanian from Thracian, another lost ancient language from farther east than Illyrian, or from Daco-Mysian, a hypothetical mixture or ancestor of Thracian, Illyrian, and the nearly unknown language of Dacia (a nearby Roman province). (We will discuss these languages in more detail in the next chapter.)

The region containing modern-day Albania became Roman territory in the second century BC. The influence of Latin on the vocabulary of Albanian was tremendous. In later times, Bulgarian provided many loanwords, and when Albania became part of the Ottoman Empire in the fifteenth century, a phase of Turkish influence was ushered in that lasted until independence was achieved in 1912. The massive overlaying of foreign vocabulary and concomitant loss of much of the native lexicon has made the phonological and morphological history of the language unusually difficult to ascertain. Thanks to the many Latin loanwords, though, we are well informed about the post-Roman phonological changes; but much had already



Map 19.1 Geographical distribution of Gcg and Tosk

happened before then that is harder to shed light on. Albanian was in fact not universally recognized as Indo-European until the second half of the nineteenth century.

19.3. Under Turkish rule, the publication of Albanian writings was forbidden; most of the history of Albanian literature has been a history of its suppression. However, Roman Catholic missionaries working mostly in the Geg-speaking area gained permission in the fifteenth century to publish religious works. The first recorded piece of Albanian is a baptismal formula from 1462; two other fragments survive from later in the same century. A small number of Christian books were produced in the following two centuries, the earliest preserved being the *Meshari* or *Missal* of Gjon Buzuku from 1555, written in Old Geg. Tosk literature originated in Albanian-speaking regions of Sicily; the first known book in Tosk dates to 1592. But we know there were earlier works which have vanished without a trace: the existence of written Albanian is already mentioned in a letter of 1332, and the first preserved books in both Geg and Tosk share features of spelling that indicate some kind of common literary language had already developed.

The language up through the eighteenth century is often referred to as Old Albanian. The modern period is dated to the nineteenth century and beyond; around 1870, with the rise of nationalistic sentiment, greater numbers of works were written in Albanian, but largely by exiles. Only after Albanian independence did Albanian literature really begin to flourish.

Among the Albanian literary monuments that are of interest to Indo-Europeanists is the northern Albanian law code traditionally attributed (most likely incorrectly) to Lekë Dukagjini (1410–81), a comrade-in-arms of Gjergj Skanderbeg, the Albanian national hero who fought against the Ottoman Turks. This law code, or *kanun* ('canon'), is derived in large part from the customary law of northern Albania, and reflects many legal practices of great antiquity. The laws governing such matters as hospitality, the rights of heads of households, marriage, blood-feuds, and payment of damages find precise echoes in Vedic India and ancient Greece and Rome.

19.4. The most notable phonological differences between Geg and Tosk are the change of intervocalic *n* to *r* in Tosk (e.g. *vera* 'the wine' vs. Geg *vena*) and the presence of nasal vowels in Geg from old sequences of vowel plus final *-n* (e.g. *shtësë* 'animal', Old Geg *shtanse*, vs. Tosk *shtazë*; the circumflex indicates nasalization). There are also morphosyntactic differences, such as in the formation of the future tense (Geg *kam me shkue*, lit. 'I have to go', vs. Tosk *do të shkoy*, lit. 'I want to go'). The dialectal split into Geg and Tosk happened sometime after the region became Christianized in the fourth century AD: Christian Latin loanwords show Tosk rhotacism, such as Tosk *murgu* 'monk' (Geg *mungu*) from Lat. *monachus*.

19.5. Albanian famously shares several morphosyntactic features with neighboring languages of the Balkans – Greek, Bulgarian, and Serbo-Croatian. All four languages form what is often called the Balkan speech area or *Sprachbund* (to use the technical term that has been imported from German). Among the features common to most or all of these languages are a suffixed definite article; the lack of a true infinitive; the use of a periphrastic construction meaning 'want to...' to express the future tense; and the use of a particle with subjunctive verb forms. Since these features were not part of the early histories of any of these languages, how they developed and diffused throughout the region is an unsettled and interesting problem.

From PIE to Albanian

Phonology

19.6. Aside from the voiceless stops *p t k* and voiced stops *b d g*, modern Albanian has a rich complement of postalveolar and palatal consonants: the alveolar affricates *c* and *x* (roughly equal to English *ts* and *dz*), the alveopalatal affricates *ç* and *xh* (like English *ch* and *j*), and the palatal affricates *q* and *gj* (similar to the preceding but pronounced farther back in the mouth). It has nine fricatives, voiceless *f th s sh* *h* (all pronounced as in English) and voiced *v dh z zh* (where *dh* is [ð], the initial sound of English *the*). The nasals are *m, n*, and palatalized *nj* (like Spanish *ñ*), and the liquids *r* (a tap), *rr* (a trill), *l*, and *ll* (a "dark" *l* like the *l* in Eng. *ball*). Finally, it has the glide *j* (like Eng. *y*). The vowel system contains the ordinary *i e a o u* as well as *y* (like German *ü*) and *ë*, a schwa-like reduced vowel.

All forms below are given in the Tosk standard unless otherwise indicated.

CONSONANTS

19.7. Stops. PIE **plain voiceless** **p* and **t* are typically preserved: *poqa* 'I cooked' < **pek^w*; *ter* 'I dry' (verb) < **tors-ejē*; *gjalpë* 'butter' < **solp-* (cp. Eng. *salve*); and *natë* 'night' < **nok^wt-*. They were lost in some consonant clusters, as in *gjumë* 'sleep' < **supno-* and *shtatë* 'seven' < **septm-ti-*. Special outcomes are seen in clusters consisting of two dentals, as in *besë* 'truce' < **bhidh-tā* (root **bheidh-* 'trust').

19.8. The voiced aspirates lost their aspiration and fell together with the plain voiced stops, as evidenced by *besë* 'truce' < **bhidh-tā* alongside *gjerb* 'I eat, drink' < **sorb-ejē-* (cp. Lat. *sorbēre* 'to absorb'); *ndez* 'I kindle' < **en-dhog^wh-ejē-* (root **dheg^wh-* 'burn') alongside *dī* 'she-goat' < **deigh-ā* (cp. German *Ziege* 'goat'); and *gardh* 'hedge, fence' < **ghor-dho-* alongside *ligë* 'disease' < **leig-* (cp. Gk. *loigós* 'ruin, death'). Word-initial **d(h)-* sometimes becomes *dh* [ð], as in *dha* '(s)he gave', ultimately from **deh₃-*.

19.9. The Albanian reflexes of the velars present special challenges. Most specialists regard Albanian as a satem language since in the vast majority of cases the plain velars and the labiovelars fell together; but some believe that it has distinct reflexes of all three velar series. The latter view is not supported by much evidence and will not be followed here. The voiceless velar *k* comes both from **k*, as in *kam* 'I have' (< **kap-mi*), and from **k^w*, as in *kë* 'whom' < **k^wom* and Old Alb. *ulk* 'wolf' (modern *ujk*) < **u^hlk^wo-*. Similarly, *g* is the outcome of **g(h)* and **g^w(h)* in such words as *ligë* 'disease' < **leig-*, *ngrenë* 'eaten' < **en-g^wrh₃-* (root **g^werh₃-* 'devour'), and *djeg* 'burn' < **dheg^wh-*. Later, *k* and *g* were palatalized to sibilants or affricates before front vowels. Sibilants are seen in such words as *pesë* 'five' < **penk^we*; *sy* 'eye' < **(o)k^wie* or **(o)k^wi* (< PIE dual **h₃(e)k^wih₁* 'both eyes'); and *zjarm* 'fire' < **g^whermo-*. Affricates are seen in *qe*, pl. of *ka* 'cow', and *shtigje*, pl. of *shteg* 'path'.

The PIE palatal velars, by contrast, became dentals, a development interestingly reminiscent of Old Persian (see §11.30), as well as Castilian Spanish. The voiceless palatal **k̑* became *th*, as in *them* 'I say' < **keh₁-mi* (cp. OPers. *θatīy* 'says') and *athët* 'sour' < **h₂ek̑-* 'sharp'. The voiced palatal **g̑(h)* became *d* or *dh*, as in *dhëmb* 'tooth' < **gombh-o-*, *mb-ledh* 'I gather' < **leg-* (*mb-* from *mbë-* 'around'), and *dimër* 'winter' < **gheimon-*. In clusters the outcome was *gj*, as in *gju* 'knee' < **g^wnu-* (cp. Eng. *knee*).

19.10. Sibilant **s*. The sibilant **s* had no fewer than five outcomes: *sh*, *th*, *gj*, *h*, and loss. No systematic pattern can be detected in their distribution that accounts for all cases; we are surely seeing the results of ancient dialect mixture. The most frequent outcome seems to be *sh*, as in *shoh* 'see', ultimately from **sek^w-* (> Eng. *see*); this is the usual outcome before another consonant, as in *shteg* 'path' above and *është* 'is' < **en-h₂esti*. This is also the fate of *s* in loanwords from Greek and Latin: *preshe* 'leek' < Gk. *práson*, *shkallë* 'ladder' < Lat. *scala*. The outcome *th* is seen in *thi* 'pig' < **sūs*, but an almost identical preform (**sū-*, from **suh₂-*) gives *shi* 'rain' (cp. Gk. *húei* 'it rains'). The curious outcome *gj* appears for example in *gjarpër* 'snake' < **serp-en-* and *gjumë* 'sleep' < **sup-no-*. Note also *gjashtë* 'six' < **s(ū)eks-ti-*, with both *gj* and *sh* in different positions in the same word (apparently one of the ancestral dialects of Albanian had at least two conditioned outcomes of **s*). In *he(l)q* 'pull, tear', it became *h* (< **selk-*), which is also the outcome of the cluster **sk*, as in *hie* 'shade' < **ski-eh₂* (cp. Gk. *skiá*) and in various verbs that once had the thematic suffix **-ske-*, e.g. *ngroh* 'make warm' < **en-g^whor-ē-ske-* (for the

suffix compare Lat. *cal-escere* 'become warm'). The sibilant disappeared word-finally, as in *thi* 'pig' above and in the rhyming word *mu* 'mouse' < **mūs*.

19.11. Resonants. The liquids and nasals remained largely intact: *gardh* 'hedge, fence' < **ghor-dho-*, *miel* 'milk' < **h₂melǵ-*, *mi* 'mouse' < **mūs*, *na* 'we' < **nos*. As noted above, Albanian has two kinds of each liquid: a tapped *r* and a trilled *rr*, as well as a light *l* and dark *ll*. Their histories are not fully clear, but trilled *rr* is typically the outcome of certain clusters, as *rrunjë* 'yearling lamb' < **uren-* and *ferr* 'hell' < Lat. *infernum*, and *ll* the outcome of single **l* between vowels, as in Geg *llanë* 'lower arm' < **ōlenā* 'elbow' and *Llezhdër* (personal name) < Gk. *Aléksandros*.

19.12. The syllabic liquids developed into liquid plus *e*, which under some conditions (§19.15) became *i*: *drekë* 'lunch' < **dr̥kʷo-* (cp. Gk. *dórpon* 'evening meal'); *drithë* 'grain' < **ghr̥zdo-*. The syllabic nasals simply became *e*: *emër* 'name' < pre-Alb. **n-men-*; Geg *dhetë* 'ten' < **dek̥n-ti-*.

19.13. The glide **j* became *gj*, as in *ngjesh* 'I gird' < **en-jōs-e-* (from **jōs-* 'gird', cp. Gk. *zōnē* 'belt' < **jōs-nā*). The glide **u* became *v*, as in *vesh* 'I dress' < **uos-eje-*.

19.14. Laryngeals. Albanian has few if any direct reflexes of the laryngeals. A proposal that some word-initial laryngeals developed into Alb. *h-* has not met with widespread approval; the putative examples all admit of other explanations. The root aorist *dha* '(s)he gave', *dham* 'we gave' is apparently from zero-grade **dh₃-t*, **dh₃-me*, in which case it preserves a reflex of a vocalized laryngeal.

VOWELS

19.15. Short vowels. The short high vowels **i* and **u* are preserved e.g. in *gji* 'bosom' < **sinos* (cp. Lat. *sinus* 'fold'), *lidh* 'I bind' < **liǵ-* (cp. Lat. *ligāre* 'to bind'), and *gjumë* 'sleep' < **sup-no-*. However, **i* was lowered to *e* if the next syllable originally contained **ā*, as in many feminine nouns, e.g. *degë* 'branch' < **duǵhā-* (cp. Eng. *twig*), *besë* 'truce' < **bhidh-tā*. Short **e* remained intact in such words as *mbledh* 'I gather' (**h₂mbbh(i)-leg-*) and *pesë* 'five' (**penkʷe*), but it often became *ie* or *je* or *ja*, as in *bie(r)* 'I carry' < **bher-*, *pjerdh* 'I fart' < **perd-e-*, *njeri* 'man' < **h₂ner-*, and *jam* 'I am' < **h₂esmi*. Short **a* and **o* merged as *a*: *kam* 'I have' < **kap-mi*, *na* 'we' < **nos*, *natë* 'night' < **nokʷt-*.

19.16. Short vowels were subject to syncope (deletion) when unstressed, as in the many verbs beginning with *n-* from the prefix **en-* 'in', e.g. *ndez*, *ngroh*, *ngrenë* above; note in *ngroh* that the original first two syllables lost their vowel (**en-gʰor-ē-ske-*). This deletion of unstressed vowels happened after the Latin loanwords entered the language, as evidenced e.g. by *mbret* 'king' < Lat. *imperator*. There was also widespread loss of final unstressed syllables. Original distinctions in stress are reflected by pairs such as *nip* 'grandson' < **népō(t)* alongside *mbesë* 'granddaughter' < **nepōt-iā* (remade from PIE **nept-ih₂*).

19.17. Long vowels. The long vowels underwent more unusual changes than their short counterparts. Long **ī* and **ū* fell together as *i*, as reflected in *pi* 'drink' < **pī-* (< earlier **pih₃-*) and *mi* 'mouse' < **mūs*. In Latin loans, *ū* was fronted to *y* (pronounced like German *ü*), as in *gymtyrë* 'joint (of the body)' < Lat. *iunctūra*. Long **ē* and **ā* fell together as *o* when stressed, as in *motër* 'sister' < **māter-* (the 'mother' word in PIE; sense development probably via 'older sister who takes a motherly role' or the like) and *plotë* 'full' < **plē-to-* (< **pleh₂-to-*). When unstressed, though, **ā* became *ë*, as in *mbesë* 'granddaughter' in the previous section. Long

**ō* became *e*, as in *resh* 'rain' < **rōs-* (cp. Lat. *rōs* 'dew') and *ne* 'us' (accusative) < **nos*.

19.18. Diphthongs. The diphthongs **ei* and **oi* became *i* and *e*, respectively: *dimër* 'winter' < **gheimon-*, *shteg* 'path' < **stoigh-o-*. The diphthong **eu* probably became *e* as well, as in *pljesht* 'flea' < **pleus-* (cp. Eng. *flea*), while **au* became *a*, as in *thanj* 'I dry' < **saus-njo-*. Sure examples of reflexes of **ai* or **ou* are unknown.

Morphology and syntax

19.19. For all the change that time has wrought on Albanian phonology, it has left the language's morphology less altered. Nouns are inflected in five cases: nominative, genitive, dative, accusative, and ablative. (A sixth case, a locative, is found in some dialects and in Old Albanian.) Both numbers distinguish an indefinite from a definite form, the latter having a suffixed definite article, with or without a separate definite article before the noun. Thus contrast indefinite *mal* 'mountain' with definite *mali* 'the mountain'; indefinite *nip* 'nephew' with definite *i nipi* 'the nephew'. The definite singular frequently preserves phonetic material that was lost in the indefinite. Thus Geg *gjëni* 'the knee' preserves the second -*n-* of the preform **gñu-no-*, lost in the indefinite sing. *gju*. Nouns come in masculine and feminine genders, but a sizeable group are masculine in the singular but feminine in the plural, as *ujë* 'water' (masc.), pl. *ujëra* 'waters' (fem.); these are old neuters.

19.20. Most of the recognizable case-endings in nouns have disappeared, as can be seen from the following sample paradigm of the masculine noun *mal* 'mountain':

	indef. sing.	def. sing.	indef. pl.	def. pl.
Nom.	<i>mal</i>	<i>mali</i>	<i>male</i>	<i>malet</i>
Gen.	<i>mali</i>	<i>malit</i>	<i>maleve</i>	<i>malevet</i>
Dat.	<i>mali</i>	<i>malit</i>	<i>maleve</i>	<i>malevet</i>
Acc.	<i>mal</i>	<i>malin</i>	<i>male</i>	<i>malet</i>
Abl.	<i>mali</i>	<i>malit</i>	<i>malesh</i>	<i>malevet</i>

The nominative plural ending -*e* continues **-oi*, the pronominal ending that replaced the original nominal ending in several other branches as well (§6.53). The postposed definite article, ultimately from the demonstrative pronominal stem **to-* (§7.9), is a feature shared with other languages of the Balkans (§19.5 above). The dative and genitive have the same endings in nouns, but are distinct in pronouns. The ablative plural ending -*sh* continues the PIE locative plural **-su* (or **-si*, if Albanian innovated in the same way as Greek; §6.17).

The definite forms come from old combinations of the indefinite and a suffixed definite article in **so-* or **to-* (the latter, for example, in *mali-t*, *maleve-t*).

19.21. The verbal system is complex; here the language has innovated more heavily. The present and aorist have been inherited from PIE, in addition to which an imperfect has been newly created. Five other tenses – perfect, two pluperfects, future, and future perfect – are formed periphrastically, that is, with a helping verb. Most remarkable is the complex system of moods: in addition to the familiar indicative, subjunctive, optative, and imperative, there is a separate conditional, as well as a mood called the admirative that is used to express surprise, wonder, or any

action not related on the speaker's own authority (as often in indirect discourse). There are two participles (present and perfect, the latter formed in part to the old IE perfect stem, which has otherwise disappeared), two infinitives (present and perfect), and two gerunds (also present and perfect). Albanian has also preserved the IE distinction in voice between active and mediopassive.

19.22. Many verbs distinguish among tenses, or among persons within a single tense, by means of vowel or consonant alternations. Thus contrast the presents *bredh* 'I run', *mbjell* 'I sow', and *vë* 'I put' with the aorists *brodha*, *mbolla*, *vura*; and contrast present 1st sing. *marr* 'I take' with 2nd sing. *merr* 'you take'. These alternations often continue PIE ablaut: thus the -o- of aorists like *brodha* and *mbolla* ultimately goes back to the *-ē- of past tenses of the type seen also in e.g. Lat. *lēgī* 'I gathered, read' (a perfect in Latin). (In fact, *lēgī* is exactly cognate with the Albanian aorist -*lodha* in *mblodha* 'I gathered'; see §5.46.) Consonant alternations are of more recent vintage; examples include *pek* 'I bake' ~ aorist *poqa* ~ present mediopassive *piqem* 'I am baked', and *djeg* ~ *dogja* ~ *digjem*, the same forms of the verb meaning 'burn'.

19.23. Only three athematic presents survive: *jam* 'I am', *kam* 'I have', and *them* 'I say', from **h₂es-mi*, **kap-mi*, and **keh₁-mi*, respectively. The imperfect in part continues the PIE imperfect, e.g. *jeshë* 'I was' < **h₂es-ṃ* (vs. *jam* 'I am' < **h₂es-mi*); but the plural endings, characterized by the vowel -i-, come from the optative (see below). The aorist continues several PIE formations. We have already seen the root aorist *dha* '(s)he gave' above in §19.14. Albanian also has some inherited thematic aorists with zero-grade of the root, such as Old Geg *kle* '(s)he was' < **k^wl-e-* (cp. Gk. *é-pl-eto* '(s)he turned'). The s-aorist is traditionally seen as reflected in such forms as *dha-shë* 'I gave' and *ra-shë* 'I fell', but in Old Albanian these forms lacked the -sh- and in PIE these roots did not form s-aorists. The source of the -sh- in these and some other forms is not entirely clear.

19.24. To illustrate verb conjugation, below are the present, imperfect, and aorist indicative and present subjunctive of the verb *them* 'I say', plus the present optative of the verb meaning 'have' (*kam* 'I have', the present, is formed from a different stem). In the rightmost column is given a sample mediopassive present, *duhem* 'I am wanted':

	pres. ind.	impf. ind.	aor. ind.	pres. subj.	pres. opt.	pres. ind. mediopass.
sg.						
1	them	thosha	thashë	them	paça	duhem
2	thua	thoshe	the	thuash	paç	duhesh
3	thotë	thoshte	tha	thotë	pastë	duhet
pl.						
1	themi	thoshim	thamë	themi	paçim	duhemi
2	thoni	thoshit	thatë	thoni	paçi	duheni
3	thonë	thoshin	thanë	thonë	paçin	duhen

Forms like *them* and *thotë* contain recognizable reflexes of the PIE 1st and 3rd sing. personal endings **-mi* and **-ti*. Even clearer are the plural paradigms *thoshim* -*it* -*in* and *thamë* -*të* -*në*, reflecting PIE **-me-* **-te-* **-nt(i)*. The mediopassive is of the

type seen in Greek (-*mai -sai -tai*) rather than Latin (-*mur -ris -tur*), i.e., the middle marker *-r has been replaced by the primary active marker *-i. The zero-grade of the PIE optative marker *-ih₂- is continued by the stem vowel -i- in *paçim -i -in*, and indirectly by the palatalization in the singular forms *paça* and *paç*.

SYNTAX

19.25. Because of its numerous inflections, Albanian word order is relatively free. In the text below, for example, the order SVO in *qi ban farë* ‘which makes seed’ exists right alongside SOV in *qi farë kâ* ‘which has seed’. Albanian, like PIE, had both stressed and unstressed (clitic) forms of the personal pronouns; characteristic of Albanian syntax is the use of the latter to double an object noun (e.g. *Zefi e mori librin* lit. ‘Joseph it took the book’) and to double a stressed pronoun preceding a verb (e.g. *mua me merr malli* lit. ‘me me takes longing’ = ‘longing takes me’). Interestingly, these enclitic pronouns, though typically not clause-initial, can appear clause-initially in commands, as *Ma ep librin* ‘Give me the book’. None of the enclitic pronouns in other IE languages can be so placed.

Old Albanian (Geg) text sample

19.26. From the *Meshari* (*Missal*) of Gjon Buzuku, the earliest surviving Albanian book, from 1555. The language is Old Geg. The passage below is a translation of Genesis 1:29–30. The spelling has been modernized, and long vowels are indicated with macrons.

E tha Zotynë: Hinje, se u u kam dhanë gjithë bārr qi ban farë për-ëmbi gjithë dhët, e gjithë pemë qi farë kâ por-si e siadó farë, juve me u klenë për gjellë; (30) e gjithë shtanse te dheut e gjithë qish fluturón për qiell e gjithë shtanse qi anshtë për-ëmbi dhët, qi shpirt i gjallë anshtë ëndë to, për-se juve t-u jenë për gjellë.

And God said, “Behold, for I have given you every plant that makes seed (which is) upon all the earth, and every fruit which has seed as well as every seed, to be food for you; (30) and to every beast of the earth, and to everything that flies in the sky and to every beast that is upon the earth that has living breath in it, in order that they be food for you.”

19.26a. Notes. E: ‘and’. tha: ‘spoke’, 3rd sing. aor. < **kēh₂*- (see §19.9). Zotynë: ‘our God’, contracted from *Zoti ynë*. The word for ‘god’ is conjectured to come from a compound **desjas-poti*- ‘master of the servants in a household’, and *ynë* ‘our’ is contracted from **a-y ne*, literally ‘the one which (is) ours’, with **ne* the true old 1st pl. pronoun, from earlier **nōs*. Hinje: ‘Behold’. se: ‘that, for’, conjunction. u u kam dhanë: ‘I have given you’. The first *u* is an old (and still dialectal) short form of *unë* ‘I’, while the second is the 2nd pl. enclitic dative ‘to you’. There follows a periphrastic perfect, formed of *kam* ‘I have’ (< **kap-mi*) plus the past participle *dhanë* ‘given’. gjithë: ‘every, all’. bārr: ‘grass, plant’, accus. sing. of *barrë*. qi: ‘who’, relative pronoun; Tosk *që*, from PIE root **k^o*-. ban: ‘makes’, 3rd sing. present of Geg *bâj* (Tosk *běj*) ‘make’. farë: ‘seed’. për-ëmbi: ‘over’, modern Tosk *përmbi*; consists of the prepositions *për* ‘to, with’ and *mbi* (older *ëmbi*) ‘upon’, cognate with Lat. *ambi*- ‘around, both’. dhët: ‘earth’, definite genit. sing. of *dhe* < pre-Albanian **ghō*, ultimately from PIE **dh(e)ghōm*; cp. Gk. *kthōn*. pemë: ‘fruit’, borrowing from Lat. *pōmum*. kâ: ‘has’, 3rd sing. of *kam* above. por-si ē lit. ‘just as also’. siadó: ‘every’, modern *cilado*. juve me u klenë:

'to you to be'; *u* is the obligatory clitic double of *juve* 'to you' (§19.25), while *me klenë* 'to be' is the infinitive (modern standard *për qenë*). The *-ve* of *juve* continues the PIE pronominal genit. pl. **-sōm* (§7.8); when **-s-* disappeared between vowels, a *-v-* developed as a transition. *për*: 'for', *gjellë*: 'meal', borrowed from Serbo-Croatian *jelo*. *shtanse*: 'animal', dat. sing., modern standard (nomin.) *shtazë*, Geg *shtasë*. *të dheut*: 'of the earth', definite genit. sing. of *dhe* 'earth' above, with preposed article *të* as well as postposed *-t*. *gjithë qish*: 'everything (that)', modern *gjithëka*. *fluturón*: 'flies', 3rd sing. present; borrowing from Vulgar Lat. **fluctulare*. *qiell*: 'heaven', borrowing from Lat. *caelum*. *anshtë*: 'is', modern Tosk *është*; the nasal in Geg (modern *është*) betrays that this is not simply from PIE **h₁esti* 'is' but from a compound **en-h₁esti*, lit. 'is in', which eventually became the ordinary way of saying 'is' in Albanian. *shpirt*: 'breath, spirit', borrowing from Lat. *spīritus*. *i*: a connective particle used in attributive adjective constructions. *gjallë*: 'living', from PIE **sol(H)yo-* 'whole' (> Lat. *salvus* 'healthy'). *ëndë to*: 'in that'. *për-se*: 'in order that', modern standard *që të*. *jenë*: '(they) be', 3rd pl. present subjunctive of *jam* 'I am'.

For Further Reading

Because of its marginal status in IE studies, Albanian has been chronically underserved as far as good reference works go; but that situation may be changing, as several appeared just in the 1990s. The book-length treatments in English are unfortunately problematic. Huld 1984, treats phonology and contains a 250-word etymological glossary of "core" words; it has some usefulness, but the reconstructions are idiosyncratic. The more recent Orel 2000 regrettably ignores Old Albanian for the most part, and also suffers from out-of-date PIE reconstructions. These features also mar Orel 1998, the only available etymological dictionary in English; it is more useful than the companion historical grammar, but omits loanwords from after the eleventh or twelfth century. For filling some of these gaps, Meyer 1891, though old, is still useful. An excellent historical grammar is Demiraj 1993; its treatment of the developments from PIE to Albanian is a bit spartan, but the other historical discussions are thorough and illuminating. Also very good is Sanz Ledesma 1996, a small but very informative Spanish overview of Albanian historical grammar. A much better treatment of the historical phonology than Huld's can be found in Demiraj 1997, which also has a full etymological glossary of inherited lexemes. One can profitably read the scores upon scores of short articles on Albanian by the American Indo-Europeanist Eric P. Hamp, such as Hamp 1977. Camaj 1984 is the best reference grammar written in English and doubles as a textbook, although it is very challenging to use; it includes facts from the non-standard dialects and from Old Albanian. Fox 1989 contains the famous law code of Lekë Dukagjini with a translation and comparative notes.

Exercises

1 Provide the Albanian outcome or outcomes of the following PIE sounds:

a	*s	c	*k	e	*gʰh	g	*d
b	*kʷ	d	*bh	f	*p	h	*ē

2 Explain the history or significance of the following Albanian forms. The forms are in Tosk unless otherwise noted:

a	murgu	e	natë	i	Geg gjûni
b	pesë	f	mbret	j	mblodha

c	ngroh	g	mi	k	Old Geg	kle
d	besë	h	ne	l		them

- How do we know that the *-e-* of Geg *dhetë* 'ten' continues the nasal rather than the *-e-* of PIE **deǵmti-*? (Assume an intermediate stage **deǵmti*.)
- In the text sample, the form *juve* 'to you' (dat. pl.) was encountered. The 2nd pl. (nomin.) *ju* looks like it should continue PIE **jūs*. Why is that derivation problematic?
- In the text sample, the 1st sing. pronoun *u* 'I' was encountered. This is normally taken to be from earlier **udh*, from PIE **eǵ-* 'I', with a replacement of **e* by *u*. In light of §7.1 and Hittite *ūk* 'I', also irregularly from **eǵ-*, provide a possible explanation for the *u*. (Don't assume any early connection between Albanian and Hittite!)
- Below is a list of some colloquial Latin loanwords (with length marks removed) that were borrowed at an early date into Albanian, with their Albanian outcomes. Describe the fate of unaccented word-initial syllables.

Latin	Albanian	gloss	Latin	Albanian	gloss
<i>parēntis</i>	<i>përind</i>	'parent'	<i>camīsia</i>	<i>këmishë</i>	'shirt'
<i>erīca</i>	<i>riq</i>	'heath'	<i>invītāre</i>	<i>ftoj</i>	'invite'
<i>Aprīlis</i>	<i>prill</i>	'April'	<i>amicus</i>	<i>mik</i>	'friend'
<i>infērnū</i>	<i>fer</i>	'hell'	<i>arēna</i>	<i>rënë</i>	'sand'
<i>Antōnius</i>	<i>Ndue</i>	'Antony'	<i>sanitātem</i>	<i>shëndet</i>	'health'
<i>Augūstus</i>	<i>gusht</i>	'August'	<i>aerāmen</i>	<i>rem</i>	'copper'

- It was stated in §19.19 that Albanian nouns that are masculine in the singular and feminine in the plural descend from old neuters. Sketch how this development might have taken place.

PIE Vocabulary XI: Utterance

- *yekʷ-* 'say': Ved. *vācas-* 'word', Gk. (*w*)*épos* 'speech', Lat. *uōx* (*uōc-*) 'voice'
**yer(h)-* 'speak': Gk. *rhētōr* 'public speaker', Lat. *uerbum* 'word', Eng. *word*
**prek-* 'ask': Ved. *prcchāti* 'asks', Lat. *precor* 'I entreat', German *fragen*, Toch. B *prek-*
**h₁neh₂mṇ* or **h₁nomṇ* 'NAME': Hitt. *lāman-*, Ved. *nāma*, Gk. *ónoma*, Lat. *nōmen*
**kan-* 'sing': Lat. *canō* 'I sing', OIr. *canim* 'I sing', Eng. *HEN*
**ǵheuH-* 'invoke (a deity)': Ved. *hāvate* 'invokes', OCS *zovǫ* 'I call'

20 Fragmentary Languages

Introduction

20.1. There remain to be discussed a half-dozen poorly understood IE languages whose preservation is so fragmentary that their exact position within the family tree is uncertain: Phrygian, Thracian, Macedonian, Illyrian, Venetic, Messapic, and Lusitanian. They are mostly known from short inscriptions, personal names, place-names, and glosses from the first millennium BC. Our survey will proceed geographically from east to west.

Since so little is known about most of these languages – especially Thracian, Macedonian, and Illyrian – the sheer variety and number of attempts at their interpretation have known no bounds. It can almost be stated as a rule in historical linguistics that the fewer the attested forms in a language and the poorer their preservation, the greater the quantity of speculative ink spilt over them. The scholarship on these languages, especially on the least well-preserved, has all too frequently thrown caution to the winds and built vast interpretive edifices with little basis in established – or establishable – fact. It is healthy to adopt a skeptical attitude with regard to much of it. Nonetheless, as we will see below, there does exist some clear (and at times valuable) IE material scattered throughout the remains of these languages.

20.2. Most of the theories concerning the interrelationships of the poorer-known languages are based on the so-called etymological method: trying to etymologize the linguistic remains that we have in light of known forms in other languages. Given that the bulk of the remains consists of personal and place-names, to which no “meaning” can be assigned in the usual sense, this method falls victim to some obvious dangers, not the least of which is circularity. Though often quite compelling on the surface, these etymologies cannot be tested against anything; they are almost always the product of seeing the morphemes that one wants to see. For example, the Thracian name *Diazenis* has been interpreted as the equivalent of Gk. *Dio-génēs* ‘born of Zeus’, and *Dizazenis* as the equivalent of *Theo-genēs* ‘born of god’ (*theo-* < Proto-Greek **theho-* < earlier **theso-*). But aside from the reasonable claim that these names contain three identifiable morphemes, *dia-*, *diza-*, and *-zenis*, we can really say nothing more with certainty. Divine names have the same problems. A well-known interpretation of *Sabazios*, the name of the Phrygian and Thracian Bacchus, is a good example: since Bacchus’s native Latin name (*Liber*) means ‘free’, it has been suggested that *Sabazios* is cognate with OCS *svobodī* ‘free’ (< **syo-bhodbhio-*). This idea is as clever and imaginative as it is speculative and unprovable.



Map 20.1 The ancient Mediterranean, showing regions and languages in this chapter

A truly embarrassing example of how the etymological method can backfire is the following. In the early 1900s a ring was unearthed near Shkodër in Albania with three words on it that were averred to be Illyrian; nearly identical forms were known from Messapic, and on their basis the inscription was given an interpretation: “To the goddess Oethe.” But the ring later turned out to be from the Byzantine period and written in Greek – it had been read backwards.

The point of the foregoing is not to disparage speculation and imaginative thinking; any science needs both to move forward. The problem arises when such speculations harden over time into facts in people’s minds and become the sole basis for further and far-reaching theories (such as the establishment of sound changes or linguistic filiation). In dealing with these fragmentary languages, it is important to distinguish at all times what is known for certain from what is only guesswork.

Phrygian

20.3. Phrygian occupies a special place in linguists’ hearts because of a famous legend related by the Greek historian Herodotus. He tells how the Egyptian pharaoh Psammetichus set about to determine what the oldest language in the world was. Accordingly he isolated two newborn infants from exposure to human speech, awaiting the day that the children would say their first words. The pharaoh reasoned that whatever language they first spoke in would be mankind’s original, oldest language. When that day came, the children’s first utterance was *bekos*, Phrygian for ‘bread’; so Psammetichus concluded that Phrygian was the world’s oldest language.

The Phrygians were a people who first enter history as immigrants into northwest Anatolia sometime after the fall of the Hittite Empire in the early twelfth century BC. They emigrated probably from the Balkans; Herodotus names Macedonia as their old homeland. It has been suggested that they were related to the Thracians (see §20.10 below) as well as to the Armenians, whose ancestors most likely came into Anatolia around the same time (see §16.1). It was formerly thought that the Phrygians’ entrance into Anatolia not only coincided with the demise of the Hittites

but also may have been a cause of it. Recent scholarship, however, tends to date their incursion to two, perhaps three centuries later.

When they had moved farther inland, the Phrygians established a state that joined a loose confederation (called the *Muški* in Assyrian records) with several other peoples to the east; this confederation was the main power in Anatolia until around the early eighth century BC. Later in that century, under the legendary king Midas, the center of Phrygian power shifted westward; his kingdom fell around 700 to the Cimmerians, invaders from the north who destroyed the Phrygian capital of Gordion (Gordium). Though the Phrygians lived on for many centuries more, they never again formed a centralized state.

20.4. While the ancient accounts of the Balkan origins of the Phrygians are generally followed, the first demonstrably Phrygian artifacts, dating to the mid-eighth century BC, are from central Anatolia. About 80 inscriptions constituting Old Phrygian date from this time and later, until about 450 BC; they are written in an alphabet apparently derived from an early version of the Greek alphabet used in Asia Minor. Then there is a long gap before the second layer of Phrygian, called New Phrygian, appears; it dates from the first and second centuries AD. Over a hundred New Phrygian inscriptions are known, written in the Greek alphabet; they are mostly funerary in nature and contain curses against grave-robbers. The inscriptions typically lack word-divisions, especially in the New Phrygian period.

Ancient writers, especially the Alexandrian glossator Hesychius (fl. fifth century AD?), also preserve some Phrygian glosses (that is, Phrygian words with a definition or explanation in Greek). The glosses seem to be reliable, since several of the words are also known from Phrygian inscriptions.

Sketch of Phrygian historical grammar

20.5. Probably more is known about Phrygian than about any of the other languages in this chapter, and some outlines of its history are clear. The PIE voiced aspirates lost their aspiration: *ab-beret* 'brings to' < **ad-bher-*; *ad-daket* 'places' < **ad-dhh-* (cp. Lat. *fac-it* 'does'); *glouros* 'gold' < **ǵhlō-ro-* (cp. Gk. *khlōrós* 'yellow-green, green'). This last example shows that Phrygian is a centum language. (Potentially contradicting this is the word *zemel-* 'slave', presumably originally 'man' and cognate with Lat. *homō* and its relatives from PIE **dheǵhom-*, **dhǵhem-* 'earth'; but the *z* here could be due to later palatalization before the *e*, or a special outcome of the original "thorn" cluster.) The plain voiceless and voiced stop series remained intact, as in *mater* 'mother' < **mātēr* and *podas* 'feet' (accusative) < **pod-ns*. The labiovelars lost their labial element, as in *ke* 'and' < **kʷe*. Of special interest is the fact that, as in Greek and Armenian, word-initial laryngeals before consonants were vocalized as vowels, as in *anar* 'man' < **h₂nēr* (cp. Gk. *anēr*, Arm. *ayr*) and *onoma* 'name' (see §6.36; cp. Gk. *ónoma*, *ónuma*, Arm. *anown*).

20.6. The inscriptions allow us to tally up five cases in the noun: nominative, vocative, accusative, genitive, and dative. There may well have been others that happen not to be attested in our meager remains. We have examples of consonant stems, thematic nouns, *i*-stems, *u*-stems, and feminine *ā*-stems; in form they resemble Greek.

20.7. Phrygian verbs were conjugated in both active and middle: active *addaket* 'causes, does', middle *addaketor* (apparently with the same meaning). In the form *egeseti* 'he will do' (?) we may have an *s*-future (see §5.38), while a preterite in *-s* is

attested in such forms as *edaes* '(s)he put' and *eneparkes* (sense uncertain). These last two forms show that Phrygian, like Greek, Armenian, and Indo-Iranian, prefixed finite past-tense forms with an augment *e-* (*e-daes*, *en-e-parkes*; see §5.40). Strikingly reminiscent of Greek are perfect middle participles in *-meno-* with reduplication, such as *tetikmenos* 'accursed' (?) and (with preverb) *protuss[e]stamenan* 'set up, established'. Third-person imperatives are found in such forms as sing. *eitou* 'let him be' (probably from **h₁eitōd* 'let him go', cp. Gk. *itō* and Lat. *itō*).

20.8. Phrygian shares more features with Greek than with any other IE language, a few of which are further shared by Armenian and Indo-Iranian. On the basis of such similarities, a subgroup of IE consisting of Greek, Phrygian, Armenian, and Indo-Iranian has been suggested (as mentioned in §10.4).

Phrygian text sample

20.9. Phrygian sepulchral curse inscribed on a doorstone from central Phrygia, second or third century AD. The text is given first in its original form without word-divisions. The first two lines are understood better than the third, and are given a provisional word-division and translation. They are formulaic, found (with minor variations) on many other New Phrygian sepulchral inscriptions. Compare the Phrygian Greek formula "Whoever will do a bad thing to this tomb, let him be cursed" (*hóstiti an tōi hērōōi toutōi kakōn poiēsei hupokatáratos éstō*).

iosnisemounknoumanekakounaddaketgegreimenane
gedoutiosoutanakkeoibekosakkalostidregrouneitou
autoskeouakerokagegaritmenosashatanteutous

ios ni semoun knoumanei kakoun addaket gegreimenan e-
gedou tios outan akke oi bekos akkalos tidregroun eitou

Whoever does evil to this grave, let him bear the . . .
curse of god, and let bread (and?) water (?) become unpalatable for him.

20.9a. Notes. *ios*: 'whoever, (he) who', relative pronoun, PIE **ios*; *ni* is apparently a particle, perhaps making the relative pronoun indefinite. *semoun*: 'this', usually taken to be a dative of **so-* 'this' with *-m-* as part of the pronominal oblique stem-formant **-sm-* (see §7.8). *knoumanei*: 'to the tomb', dat. sing.; etymology disputed. *kakoun*: 'bad, evil', some sort of accusative singular showing that **-m* became *-n* in Phrygian, as in Greek. Cp. Gk. *kakōn* 'evil, an evil thing'. *addaket*: 'causes, does (to)', probably consisting of a preverb *ad-* 'to, at, upon' (cp. Lat. *ad* 'to') plus *dak-* 'do', from **dhh₁-k-*, extended zero-grade of **dheh₁-* 'place, put, do' (cp. Lat. *fa-cere* 'to do', Gk. aorist *(ē)thē-k-e* 'he did' < full grade **e-dheh₁-k-*). *gegreimenan*: feminine accus. sing. perfect mediopassive participle, in formation like that of e.g. Gk. *lelumenēn* 'having been released'. The meaning is uncertain but apparently negative; it modifies *outan*, usually understood as 'curse'. Note the poetic word order, with the participle separated from the noun it modifies by an intervening verb and dependent genitive. *egedou*: 'let him bear' or the like, perhaps from **h₂eg₂-e-tōd*, 3rd sing. imperative; compare *eitou* at the end of the line without weakening of the *-t-* (due to being directly after the stress?). *tios*: 'of god', perhaps PIE **diu₁-os*. It could also be the name of a deity ('Tis'). *akke*: usually considered to mean 'and' and compared with Latin *atque*. *oi*: 'to him', dative, cp. Gk. dat. *hoi*. *bekos*: 'bread', see discussion in §20.3. *akkalos*: 'water' (?), cp. Lat. *aqua*. If this passage is correctly interpreted, it is reminiscent of the curse of the legendary Phrygian king Midas,

who wished that everything he touched should turn to gold – the unfortunate consequence being that all the food and drink he touched he could no longer consume, *eitou*: perhaps a 3rd person imperative in **-tō(d)* like Lat. *estō* ‘let it be’. The third line of the inscription probably begins with *autos* ‘he himself’, a word also found in Greek. This line contains another middle participle, *gegaritmenos*, perhaps cognate with Gk. *kekharisménos* ‘devoted’.

Thracian

20.10. In classical times, Thrace (Thracia) was a region in the eastern Balkans, located west of the Black Sea between the Danube River to the north and the Sea of Marmara to the south, in what is now mostly Bulgaria. Herodotus described the Thracians as the greatest and most populous people on earth after the Indians; they are mentioned in Homer as allies of the Trojans during the Trojan War. Their prowess as fighters was legendary; they were also famous for their music and poetry, none of which, regrettably, has survived.

Numerous brief coin inscriptions from as early as the sixth century BC have been found in what was once Thrace, as well as a few short inscriptions that have no agreed-upon interpretation. The coin legends and various Classical authors preserve some Thracian personal and geographic names. Personal names are attested up until the sixth century AD.

20.11. Ancient glossators, especially Hesychius, preserve about eighty or ninety words that are said to be Thracian; of these, no more than three dozen can be confidently labeled as such. Here we can identify a few that are in all likelihood Indo-European, such as *briza* ‘spelt, rye’ (PIE **urgh-jā* ‘rye’, cp. German *Roggen*), *br(o)utos* ‘drink made from barley, beer’ (PIE **bhreu-* ‘bubble, ferment’, the source of Eng. *brew* and *broth*), and *-para* ‘ford, pass’ in certain place-names (PIE **per-* ‘to cross over’, the source of Eng. *ford*). A Roman inscription from AD 226 has a Thracian word *midne* ‘homestead’ that has been plausibly compared with Latv. *mitne* ‘dwelling’ and Av. *maēθana-* ‘dwelling’.

All attempts to relate Thracian to Phrygian, Illyrian, or Dacian (the language of the neighboring province of Dacia, preserved almost solely in some plant names in Hesychius) are likewise purely speculative. The notion of a “Thraco-Phrygian” branch of IE had currency for some time but has fallen out of favor. Our knowledge of these languages is simply too limited for claims of this kind; even the notion that what the ancients called “Thracian” was a single entity is unproven.

Macedonian

20.12. The Macedonians were a group of ancient tribes located north of Greece. They are thought to have come from the west and to have slowly migrated south-eastward into lowland areas around the basin of the Haliacmon (Aliákmon) river in north-central Greece, along which they set up Argeai, their ancient capital. According to legend, they were unified under one Temenides in the seventh century BC. Their territory gradually increased to encompass Illyria, Thrace, and Phrygia. Under Philip II (382–336 BC), Macedonia defeated Greece, and his son Alexander III (Alexander the Great, 356–323 BC) famously extended Macedonian conquest to much of the known world. Following his death, Macedonia was riven by internal

conflicts until the Antigonid dynasty (277–168 BC), after whose decline the country became a Roman province (146 BC).

Our knowledge of Macedonian is limited to glosses (again, mostly in Hesychius) and a number of personal names. The inscriptions found in Macedonia are in Greek, adopted as an official language by the fifth century BC. Already in antiquity, Macedonian was regarded as bearing a close affinity with Greek, an impression that the glosses confirm. However, it is debated whether Macedonian was a rather deviant Greek dialect or a separate but closely allied language. Speaking for the latter is the one rather important Macedonian sound change that can be deduced from the glosses and personal names: the voiced aspirates lost their aspiration and became plain voiced stops, unlike any known Greek dialect (Macedonian *abroutes* 'eyebrow' = Gk. *ophrūs*; *danon* 'death' [accusative] = Gk. *thánaton*; *dōrax* 'spleen' = Gk. *thōraks* 'thorax'; personal names *Ber(e)nika* = Gk. *Pherenikē* 'bearing victory' and *Bilippos* = Gk. *Philippos*). Attempts to link Macedonian with Thracian and/or Illyrian in various ways are quite inconclusive.

Illyrian

20.13. The Illyrians, while not instantly familiar to most people today, were rather important in ancient history. They or their ancestors had settled in the Balkans perhaps in the early Bronze Age in the first half of the second millennium BC, and gradually expanded and founded various local kingdoms over the ensuing centuries. They came under the influence of Greek culture between the eighth and sixth centuries BC, when the Greeks set up colonies in the area, most notably Epidamnus (Dyrrhachium to the Romans, modern Durrës in Albania). After a series of conflicts with Rome, the Illyrians became Roman dependents by 165 BC; their lands constituted the new Roman province of Illyricum. Although they were to play an important role in the Roman army and even furnished later Rome with several famous emperors (including Diocletian, Constantine the Great, and Justinian I), the Illyrians never became fully assimilated Romans and kept their language.

What that language was is uncertain. The term "Illyrian" has referred to many things, until recently to any non-Celtic language in the broad area west of Thrace, north of Greece and Macedonia, and east of the Veneti (northeastern Italy). But scholarly work beginning in the 1960s has shown that the region is neither archaeologically nor onomastically uniform and that it breaks down into three distinct cultural and linguistic areas, of which only one can properly be called Illyrian. No treatment yet exists of the linguistic remains from just this region to the exclusion of the others. The remains consist of some personal and place-names and some glosses. There are no known inscriptions (and recall §20.2 above).

20.14. Two untestable hypotheses about Illyrian's connection to other languages are widely held: that Illyrian is the same as or closely related to Messapic, and that Illyrian is the ancestor of Albanian. The first hypothesis is based on the close cultural connections between the Messapians and Illyrians, and on certain similarities between some linguistic elements. The second hypothesis has very little, if any, linguistic support, but makes geographic sense; proponents also point out that the word *Albanoí* 'Albanians' is first attested (in the *Geography* of Ptolemy) as the name of an Illyrian tribe. One glossed word that has been compared with Albanian is *rhinos* 'fog' (cp. Old Geg *ren* 'cloud', modern *rê*), but this alone does not prove

the case. The possible relationship to Messapic does not help, for the Messapic inscriptions evince no obvious similarities to Albanian.

A "Thracio-Illyrian" branch of Indo-European has been proposed by some who view Illyrian and Thracian as related. Others have proposed a mixture called "Daco-Mysian" or simply Dacian, a combination of Thracian, Illyrian, and the language of the neighboring Roman province of Dacia. All such proposals have very little to go on and are premature.

20.15. A few words of IE interest that are traditionally called Illyrian are: *Deipaturos*, the Illyrian name for Jupiter or Zeus that continues the PIE name 'Father Sky' (**d̥iēus ph₂tēr*, cp. Lat. *Iū-piter*; see §2.19); a word *teuta-* meaning 'people' in personal names from the root **teutā* found in western Europe (§2.8); the personal name *Vescleves*, usually understood to mean 'having good fame' and composed etymologically of the same elements as the Sanskrit adjective *vasu-śravas-*; and *sabaia* or *sabaium*, mentioned by the late Roman historian Ammianus Marcellinus and defined as a beer-like drink, which may be from the same root **sab-* that gives Germanic words for 'juice' (German *Saft* 'juice', Eng. *sap*). See also the notes to line 27 of the Umbrian text in §13.71a.

Venetic

20.16. Venetic was spoken in northeast Italy by the Veneti, a people famous in the Greek world for their horses. They appear to have arrived there by the end of the second millennium BC. The center of Venetic culture was the town of Este, about fifteen miles southwest of Padua; it was also the center of the cult of their main divinity, the goddess Reitia. But their territory stretched as far west as Verona and as far south as the Po valley. In the latter region they came into contact with Etruscan settlers around the sixth century BC, from whom they learned the alphabet. Venetic inscriptions, which number about 200 and date from the sixth to the first centuries BC, are written in both the Etruscan and, later, the Roman alphabet. Modern scholars sometimes divide Venetic into Archaic Venetic (550–475 BC), Old Venetic (475–300 BC), New Venetic (300–150 BC), and Veneto-Latin (150–100 BC). The inscriptions are all rather short, none being more than about ten words long; roughly half the attested words are personal names.

20.17. In view of its geographical location, it is not surprising that Venetic bears close affinities with Italic. Some of the similarities include the development of certain voiced aspirated stops to *f* (written variously *ph* or *vh*), such as the dative pl. *louzerophos* 'for the children' (< **-bhos*) and the *s*-aorist *vhagsto* 'he did' (cp. Lat. *fac-ere* 'to do', late PIE **dhh₁-k-*). However, other outcomes are seen as well that confuse the picture: alongside the dative plural ending in *-phos* is also one in *-bos* (e.g. *andeticobos* 'for the sons of Andetios' [?]), and **dh* became *z* in the word *louzerophos* seen above (from **loudhero-*).

20.18. In spite of the paucity of material, we know some outlines of the grammatical structure of the language. Five cases (nominative, accusative, genitive, dative, and instrumental) are preserved, and all three numbers – singular, dual, and plural – are found. The only known verb-forms are in the third person, such as singular present *atistei* 'sets up (?)', past *donasto* 'gave'; plural *donasan* 'they gave'. The 3rd sing. *tolar* or *toler*, of unclear meaning, is a middle form. Unlike any of the attested Italic languages, Venetic still has a living *s*-aorist, as in *dona-s-to* and *vhag-s-to* above. The language also had active *nt*-participles (e.g. *horvonte*, a masculine nominative

dual) and middle participles in *-mno-* (e.g. *alkomno* ‘?’). Of particular interest are some features of the pronominal system that find exact parallels in Germanic, and only there: a stem **selbh-* for the reflexive pronoun (Venetic *sselboisselboi* ‘for himself’, cp. the similarly doubled OHG form *selb selbo*), and a first-person accusative singular pronoun that is a rhyme formation with the nominative (Venetic *meḡo* ‘me’ rhyming with *ego* ‘I’; cp. German *mich* ‘me’ rhyming with *ich* ‘I’). Noteworthy also is an o-stem genitive singular in *-i*, matching Latin, Celtic, and Messapic.

20.19. The Venetic alphabet does not contain separate letters for the voiced stops. However, it contains a special *t*, sometimes transliterated *t²*, that stands for *d*, as in the word *t²eivos* = *deivos* ‘god’. Venetic orthography generally makes use of the Etruscan system of “syllabic punctuation,” treated in the Exercises below.

Venetic text sample

20.20. Inscription found just north of Monselice (a town east of Este and south of Padua), dating from the Archaic Venetic period, before 475 BC. There are no word-divisions in the original. The interpretation follows that of Michel Lejeune, *Manuel de la langue vénète*, where the inscription appears as number 75ter.

ego vhonei ersiniioi
vineti karis vivoi oliialekve murtuvoi atisteit

20.20a. Notes. *ego*: ‘I’, cp. Lat. *ego*. *vhonei ersiniioi*: ‘for Fons Ersinius’, a personal name in the dative. *Vhont-* is an athematic noun with dative in *-ei* (§6.9), and *ersiniioi* may be a patronymic adjective in **-jo-* (‘son of Ersinus’). *vineti karis*: uncertain; perhaps a single word. Lejeune suggests ‘friend of Vinetus’. *vivoi*: ‘alive’, dat. sing.; cp. Lat. *vivus* ‘alive’. *oliialekve*: ‘and . . .’; *-kve* is from PIE **kʷe*; *oliiale* is not understood but is perhaps an adverb. *murtuvoi*: ‘dead’; cp. Lat. *mortuus*. *atisteit*: see above, §20.18. The sense of the whole inscription is therefore something like, “I (am) for Fons Ersinius. A friend of Vinetus (?) sets up (thus?) for (him), alive and . . . dead.”

Messapic

20.21. Messapic is known from close to 300 inscriptions from southeast Italy in Calabria and Apulia and dating from the sixth to the first centuries BC. The language was spoken by an ancient people known both as the Messapii (or Messapians) and Iapyges. They are linked by ancient historians with Illyria, across the Adriatic Sea; the linkage is borne out archaeologically by similarities between Illyrian and Messapic metalwork and ceramics, and by personal names that appear in both locations. For this reason the Messapic language has often been connected by modern scholars to Illyrian; but, as noted above, we have too little Illyrian to be able to test this claim.

Messapic is written in an offshoot of an Ionic Greek alphabet that diffused early into southern Italy. Most of the inscriptions just contain personal names, and usually lack word-dividers, rendering interpretation of the longer inscriptions very difficult. A few things are known about its history and structure. Short **o* became *a*, as in the athematic genitive *damatras* ‘of Demeter’ (with genitive singular *-as* from **-os*). Like Celtic, Italic, and Venetic, it has a genitive singular ending *-i* (written *-ihi*), but unlike these other languages the ending was added to the **-o-* of o-stem nouns

rather than in place of the stem: *dazimaihi malohiaibi* 'of Dazimas Malohias'. A dative plural of the usual type is seen in *logetibas* 'for the logeti's' < *-ibhos. A few verb forms have been tentatively identified, such as *hipakaθi* 'sets up' (?), a 3rd-sing. present, and *hipades* 'set up' (?), apparently an s-aorist.

Probably the most valuable and interpretatively undisputed piece of Messapic is the formula *kl(a)ohi zis* found at the beginning of some inscriptions, which translates as "Listen, Zis (= Zeus)!" See the Notes to the text below.

Messapic text sample

20.22. Inscription from the second or first century BC from Basta (modern Vaste, a village in Apulia west of Otranto), copied down in the sixteenth century; the original has vanished. The division into lines below is probable but not assured; there are no word-divisions. We cannot translate the inscription, but we can provisionally divide it up analytically into words and structurally repeating units, following Hans Krahe, *Die Sprache der Illyrer*, vol. 1 (1964), pp. 27–8, with modifications.

klohizisθotoriamartapidovasteibasta
veinanaranindaranθoavastistaboos
šonedonasdaštassivaanetosinθitrigonošo
astaboosšonetθihidazimaihibeileihi
inθireššorišoakazareihishonetθihiotθeihioi
dazohonnihiinθivastima
daštaskraθecheihiinθiardannoapoššonnihia
imarnaihi

klohi zis	
θotoria marta pido vastei basta veinan aran	
in daranθoa	vasti staboos šonedonas
	daštas-si vaanetos
5 in-θi trigonošo	staboos šonetθihi
	dazimaihi beileihi
in-θi reššorišo	kazareihishonetθihi otθeihioi dazohonnihi
in-θi vastima	daštas kraθecheihi
in-θi ardanno	poššonnihi aimarnaihi

20.22a. Notes. *klohi zis*: 'Listen, Zis', invocation to Zis (Zeus). *Klohi*, spelled *klaohi* in older inscriptions, is probably the exact cognate of Ved. *śróṣi* 'listen', a singular s-aorist imperative. *Zis* is probably directly continues **d̥iēus*, although the phonological details are unknown. A later Greek writer attests the name *Dis* from Tarentum in southern Italy, which may well be the same thing. *pido*: supposedly *(e)pi-dō-t 'gave, presented' or the like, with preverb **epi-* (Gk. *epi-*) and a root aorist of **deh-* 'give'. *in-θi*: -θi means 'and' and likely continues PIE **k'e*. It also occurs following *daštas* in line 4, assimilated to -si, and following *otθeihioi* in line 7. Most of the words in the right-hand column are bipartite personal names in the genitive singular.

Lusitanian

20.23. In the western Iberian peninsula, between Guadiane and Duero, three inscriptions in an IE language have been found, written in Latin letters and dating

from around the time of the early Roman Empire (first century AD). This area was part of the Roman province called Lusitania, named after a people called the Lusitani; the language of these inscriptions has therefore been dubbed Lusitanian. The language is regarded by some to be Celtic; but this is quite uncertain. The identification of the language as IE is based on such forms as *doenti* 'they give', thematic dative singulars in *-oi* or, more commonly, *-ui* (the same as the corresponding Celtiberian ending), and accusative singulars in *-m* (e.g. *porcom* 'pig', *taurom* 'bull'). Note also the nominative plural *Veammicori*, a name of a people, reminiscent of *Petrucorii*, the name of a Gaulish tribe.

For Further Reading

The corpora of all these languages continue to grow; the standard collections – few of them in English – contain most of the known inscriptions but lack those that have come to light more recently. For Phrygian, the most complete and recent comprehensive collection is Orel 1997, which includes both the material from the two standard collections (Haas 1966 and Brixhe and Lejeune 1984, the latter a very fine work) and the inscriptions unearthed since then. Orel's work also contains a grammar and vocabulary, but treat the interpretations with caution. An excellent comparison of Phrygian and Greek is Neumann 1988. The Thracian texts are in Detschew 1957. For Macedonian, see Kallérís 1954–76. The standard collection of Illyrian remains (but not in the narrow sense discussed in §20.13) is Krahe 1955–64, the second volume of which contains also the Messapic inscriptions (edited by C. de Simone) and a study of Illyrian personal names by Jürgen Untermann. For a good critical review of scholarship on Thracian and Illyrian, see Katičić 1976. The Venetic inscriptions known until the mid-1970s are collected in Lejeune 1974, which also contains detailed discussions of the orthography, phonology, and morphology. The Lusitanian corpus is contained in Untermann 1997.

Exercises

- Comment on the history or significance of the following forms:

a Phryg. <i>addaket</i>	d Illyr. <i>Deipaturos</i>	g Ven. <i>sselboisselbol</i>
b Phryg. <i>anar</i>	e Illyr. <i>Vescleves</i>	h Messap. <i>-ihi</i>
c Phryg. <i>edaes</i>	f Ven. <i>vhagsto</i>	i Messap. <i>kl(a)johi zis</i>
- Phrygian *bekos* looks a lot like English *bake*, but what difficulty attends equating these etymologically?
- It was mentioned in §20.8 that many scholars believe Phrygian forms a subgroup with Greek, Armenian, and Indo-Iranian. What branches of IE do the Phrygian middles in *-tor* pattern with?
- Venetic "syllabic punctuation." Most Venetic inscriptions contain what at first appears to be a profligate fondness for interpuncts. There is, however, a fairly simple phonological principle behind their use. Interpuncts were used to frame certain kinds of sounds in certain syllabic contexts. If this framing would have led to a double interpunct, that was usually (though not always) simplified to a single one.

- a Based on the following examples (whose meanings are irrelevant for the present purposes), determine the rules for interpunct placement. Recall that *v*h stands for a single consonant (the fricative *f*).

.e.kupetari.s.	vha.g.s.to	lemeto.r.na
ka.n.te.s.	.e.go	
.o.p.po.s.	ne.r.ka	

- b What additional feature is reflected by the following?

.a.kuto.i.	.a.vhro.i.	ka.n.ta.i.
	votu.n.ke.a.	

- c Now that you have established the basic principles, what information do the preceding examples tell us about the pronunciation of the initial *i* in *iuva.n.te.i*?
 d How many syllables does the word spelled *vo.t.te.i.ilo.s.* contain? Explain.
 e Now that you have established these principles, what do the following forms tell us about Venetic syllabification? State your answer as generally as possible. (Hint: compare the remarks on Latin *patre* in §1.10.) Treat *ll* as identical to *i*.

.e.kvo.n.	ve.i.gno.i.	u.r.kli
.a.vhro.i.	iiuva.n.tii.o.	mu.s.kia.l.na.i.
vhu.g.slia	.o.s.tiala.i.	

- 5 Imagine that you are the proud discoverer of a hitherto unknown ancient IE language belonging to a hitherto unknown branch of the family. Your task is to report your discovery to the scholarly world. Describe your language, including at least the following information:

- The date of the text(s) you have found and the place of discovery;
- The outcomes of all the PIE sounds – consonants, vowels, and diphthongs – in your language. Include at least two sound changes that are conditioned, i.e., that happened only in particular phonetic environments (some of the conditioned sound changes that we've talked about are rhotacism in Latin, umlaut in Germanic, Verner's Law, and palatalization). Be sure to specify what the phonetic environments were (beginning of a word, between vowels, before a front vowel, word-finally, etc. etc.);
- The outcomes of these PIE forms: **ph₂tér* 'father', **mátēr* 'mother', **bhrátēr* 'brother', **syésōr* 'sister', **pods*, **ped-* 'foot', **mūs-* 'mouse', **k^wel-* 'to turn', **h₃erbh-* 'transfer to another sphere of ownership', **kiéyos* 'fame', **ǵ^helk^wos* 'wolf', **ǵheimōn* 'winter', **sneig^h-* 'to snow';
- A brief description of the nominal system, including: what cases are preserved; what numbers; what genders; the general fate of athematic and thematic nouns;
- A brief description of the verbal system, including: what tenses are preserved; what numbers; the general fate of athematic and thematic verbs, of the aorist, and of the perfect;
- The paradigm in the singular and 3rd plural of the descendant of **h₃es-* 'be' in the present tense, **bher-* 'carries' in the present tense, and **ǵoide* 'knows';

- g** A sample text in your language of a dozen words, including at least half that have an IE etymology and are different from the ones you give in (c) above;
- h** Some brief remarks about the culture, mythology, society, etc. of the people that spoke your language.

PIE Vocabulary XII: Basic Physical Acts

- **dheh₁*- 'place, put': Hitt. *dāi* 'puts', Ved. *dádhatī* 'puts', Gk. *títhēmi* 'I put', Lat. *faciō* 'I do', Eng. *do*
- **werǵ-* 'work': Av. *verezileiti* 'works', Gk. (*w*)*érgon* 'work', Arm. *gorc* 'work'
- **bheid-* 'split': Ved. *bhináti* 'splits', Lat. *findō* 'I split', Eng. *bite*
- **sek-* 'cut': Lat. *secō* 'I cut', Eng. *saw*, OCS *sěko* 'I cut'
- **ǵheu-* 'pour': Ved. *juhótī* 'pours', Gk. *khé(w)ō* 'I pour', Lat. *fundō* 'I pour', Eng. *ingot*

Glossary

For items not listed here, see the Index.

ablative In nouns and pronouns, the case used to express source or origin and the object of certain prepositions, especially those indicating motion away from or out of. In some languages it is also used to express instrument, accompaniment, or location.

ablaut A change in the vowel of a root or word that signals a change in grammatical function or category. The change of vowels in *sing sang sung song* is an example of ablaut.

accusative In nouns and pronouns, the case used to express the direct object of a verb and the object of certain prepositions, especially those indicating motion towards, into, across, or through.

active In verbs, the voice indicating that the grammatical subject does the action.

affricate A consonant produced by stopping the flow of air completely and then releasing the stoppage gradually, resulting in a sound that begins as a stop and ends as a fricative. English *ch* and *j* are affricates. The verb to *affricate* means to change or cause to change into an affricate.

allomorph One of the variants of a morpheme. The three pronunciations of the English plural morpheme *-(e)s* (as in *bets*, *beds*, *midges*) are allomorphs of one underlying plural morpheme.

allophone One of the pronunciations of a phoneme. The aspirated *p* of English *pit* and the unaspirated *p* in *spit* are allophones of the English phoneme [p].

apocope The loss of a vowel or syllable at the end of a word.

articulation The physical production of a speech sound.

aspirate A consonant immediately followed by a puff of breath. Sanskrit *bh* was a voiced aspirate.

assimilation The change of a speech sound such that it becomes more similar to another sound, usually a neighboring one.

case A form of a noun, pronoun, or adjective that indicates a particular grammatical role.

clitic A word that has no independent word-stress and attaches to a neighboring stressed word, with which it forms a phonological unit. The *-n't* of English *didn't* is a clitic.

cognates Words in two or more languages that are descended from the same word in their common ancestor, like Latin *pater* and English *father*.

comparative method The method by which two or more languages are systematically compared to determine whether they are descended from a common ancestor.

copula The verb *be* or its equivalent in other languages in the function of identifying the predicate with a subject, as in the sentence *Humans are mortals*.

coronal Produced with the front part (blade) of the tongue, such as *t*, *s*, *r*, and *l*.

dative In nouns and pronouns, the case used to express the indirect object, the one that benefits from an action, and sometimes possession.

denominative A verb formed from a noun, such as Eng. *to book*, *to computerize*.

- dental** Produced with the tip of the tongue touching the teeth or the ridge behind the upper teeth, as *t*.
- derivative** A word derived from another word, typically by a productive morphological process; also, a word that is historically descended from a word or word root in an ancestral language.
- derive** To form a word from another word or word root, typically by a productive morphological process.
- determiner** A modifier that acts to select a particular referent or referents. Determiners include definite and indefinite articles, demonstratives, possessive pronouns, and indefinite pronouns.
- deverbative** Formed from a verb; also, a word formed in this way.
- diachronic** Viewed or occurring over a period of time; pertaining to a language's development through time; historical. Compare *synchronic*.
- dissimilation** The change of a speech sound such that it becomes less like another sound, usually a neighboring one.
- dual** The grammatical category indicating two in number.
- enclitic** A clitic, especially one that is phonologically dependent on a preceding word.
- epenthesis** The introduction of a vowel in pronunciation, especially between two consonants or neighboring a consonant cluster, as when *athlete* is pronounced [æθəlit].
- finite** Referring to a verb form that is inflected for tense, mood, voice, person, and number.
- frequentative** A verb form indicating repeated or habitual action.
- fricative** A consonant produced by partial obstruction of the vocal tract, resulting in hissing or turbulent noise, such as *s* and *f*.
- gender** A grammatical classificational category of nouns and some pronouns. Typical gender systems involve a three-way distinction between masculine, feminine, and neuter, or a two-way distinction between animate and inanimate. Grammatical gender sometimes overlaps with biological gender.
- genitive** In nouns and pronouns, the case used to indicate possession.
- glide** A consonant, such as *w* or *y* in English, produced with relatively little obstruction of the flow of air and resembling a vowel; also called "semivowel."
- gloss** A translation. Ancient glosses, such as those in Old Irish, are typically brief interlinear or marginal comments or translations.
- glottis** The muscular ring at the upper opening of the larynx.
- guttural** Produced in the back of the mouth or in the throat; also, another term for *velar*.
- heavy** Ending in a consonant or containing a long vowel or diphthong. Said of syllables.
- high** Produced with the tongue raised toward the roof of the mouth, as the vowel sounds in the words *street* and *boot*.
- historical phonology** (morphology, syntax) The study or account of the history and development of a language's phonology, morphology, or syntax.
- imperative** In verbs, the mood used to express orders and commands.
- indicative** In verbs, the mood used to express factual statements.
- infinitive** A verbal noun, especially one used as a complement of another verb. English infinitives consist of *to* plus the verb.
- inflect** To change form, as by the addition or change of a suffix, in order to indicate change of grammatical function.
- inflection** A form of a word having a particular grammatical function, such as the accusative case of a noun or the 3rd person singular of a verb.
- instrumental** In nouns and pronouns, the case used to indicate instrument or means.
- labial** Produced with the lips, such as *p*.
- labiovelar** A velar consonant accompanied by rounding of the lips, like Proto-Indo-European **kʷ*.
- lexicon** The vocabulary of a language.
- light** Ending in a short vowel. Said of syllables.

liquid An *r* or *l*.

loanword A word borrowed from one language into another.

locative In nouns and pronouns, the case used to express location or place at, in, or on which.

manner of articulation The way in which, and the degree to which, obstruction of the vocal tract is effected during the production of a consonant, such as stoppage or frication.

medial Occurring in the middle of a word.

mediopassive Combining middle and passive voice.

middle In verbs, the voice indicating that the grammatical subject is acting upon itself or is otherwise internal to the action of the verb.

mood A grammatical category of verbs indicating the speaker's attitude or stance taken towards the action expressed by the verb.

morpheme The smallest unit of a language that has meaning. Two examples from English are the plural suffix *-s* and the word *word*.

morphology The system of rules for combining morphemes into words in a language, or the study thereof.

murmur The breathy voice characteristic of voiced aspirate consonants.

nasal A stop produced by letting air flow out through the nasal passages, such as *m* and *n*.

nominal Pertaining to a noun; having the function or properties of a noun.

nominalize To turn another part of speech into a noun, as in *the good and the bad*.

nominate In nouns and pronouns, the case used to express the subject of a sentence or a predicate nominative.

number The grammatical category indicating quantity, such as singular, dual, and plural.

obstruent A consonant that is not a resonant; a non-nasal stop, fricative, or affricate.

optative In verbs, the mood used to express wishes.

palatal Produced with the tongue touching or coming near the hard palate. The glide *y* and the *k* of the English word *kit* are palatal.

palatalize To change a consonant into a palatal consonant or into a consonant produced closer to the palate.

participle A verbal adjective. As an adjective it can modify a noun and be inflected in case, number, and gender; as a verb it can take an object and distinguish tense and voice.

passive In verbs, the voice indicating that the grammatical subject is acted upon by the action of the verb.

periphrastic Relating to a morphological category, such as a verb tense, that is expressed using an auxiliary rather than by a single inflected form, such as the English present perfect (*have gone*).

person A grammatical category of verbs and pronouns that indicates the person speaking (first person), person being spoken to (second person), or a third party (third person).

phoneme The smallest linguistic unit that can contrast meaning; a basic sound of a language.

phonology The system of sounds in a language, or the study thereof.

phonotactic Pertaining to the *phonotactics* of a language, the rules for combinations of sounds particular to a given language.

place of articulation The place in the vocal tract where the flow of air is obstructed during the production of a consonant.

possessive Another word for *genitive*.

postposition A word functionally equivalent to a preposition but occurring after the word it governs.

predicate nominative A noun, pronoun, or adjective that is predicated of a subject by means of a copula or copulative verb, such as *be*, *become*, or *seem*, and that stands in the nominative case.

preverb An adverbial particle that modifies the meaning of a verb and often appears attached to the verb as a prefix.

proclitic A clitic that is phonologically dependent on a following word.

- productive** Able to be used to form new words or inflectional forms. Said of morphemes or morphological processes, such as English *-ed*, *-ize*, *re-* (as opposed to, e.g., the *-th* of *growth*).
- pronominal** Pertaining to pronouns.
- prosody** The organization of speech sounds into rhythmic units or groups; also, the rules by which it is so organized.
- proto-language** A language not preserved in records that has been reconstructed on the basis of a known language or languages, of which it is the ancestor.
- reflex** A historical descendant or outcome of a specified ancestral sound or form. *F* is the Germanic reflex of PIE **p*, for example.
- resonant** A consonant capable of becoming the center of a syllable and thereby becoming syllabic, such as the liquids and nasals.
- rhotacize** To turn into *r*.
- sandhi** The process whereby one or more sounds at the juncture of two words or morphemes assimilate to one another or undergo other phonetic change.
- semivowel** *see* glide.
- sibilant** A fricative having a hissing quality, such as *s*.
- spirantize** To turn into a fricative, as when a *p* changes to *f* or a *b* changes to *v*.
- stem** The form to which grammatical endings are added; the form of a word minus its grammatical endings.
- stop** A consonant produced by total closure of some part of the vocal tract, whereby the flow of air is briefly stopped and then released. *P* and *d* are stops. Also called *plosive*.
- subjunctive** In verbs, a mood typically expressing non-factual modalities, as in clauses of purpose or result, conditions, and potentialities.
- substantive** A noun.
- surface** At the linguistic level of spoken language rather than at a deeper or more abstract (underlying) level.
- syllabic** Occupying the center or syllable peak of a syllable; functioning as a vowel.
- synchronic** Viewed at or characteristic of a particular point in time in the history of a language. Compare *diachronic*.
- syncope** The loss of a vowel or syllable in the interior of a word.
- syntax** The system of rules for combining words into phrases, clauses, and sentences in a language.
- tense** The grammatical category of verbs that expresses the time at which an action happens.
- umlaut** The partial assimilation of a vowel to a vowel in a neighboring (usually following) syllable.
- underlying** At a deep or abstract grammatical level rather than on the surface phonetic level.
- velar** Produced with the tongue touching the velum (soft palate, behind the hard palate). The *g* in English *got* is a velar stop.
- vocalic** Pertaining to or characteristic of vowels; also, pronounced with the addition of a vowel sound, as a vocalic laryngeal.
- vocative** In nouns and pronouns, the case used for direct address.
- voice** The verbal category expressing the role taken by the subject in the action of a verb.
- voicing** The vibration of the larynx during the production of a speech sound. *B* and *m* are voiced sounds; *p* and *b* are voiceless.
- vowel harmony** A property of some languages whereby the vowels in a given word must be alike with respect to some feature, such as height, backness, or roundness.

Bibliography

This bibliography is not intended to be even remotely exhaustive. All the titles except those marked with an asterisk are discussed in the "For Further Reading" sections at the end of the relevant chapter. Note that titles mentioned in more than one "For Further Reading" section are not listed here multiply.

Chapter 1: Introduction

- Bader, Françoise, ed. 1997. *Les Langues indo-européennes*. 2nd edition. Paris: CNRS.
- Beekes, Robert S. P. 1995. *Comparative Indo-European Linguistics: An Introduction*. Amsterdam: Benjamins.
- Brugmann, Karl. 1897–1916. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Strassburg: Trübner. (Translated as *Elements of the Comparative Grammar of the Indo-Germanic Languages* [Varanasi: Chowkhamba Sanskrit Series Office, 1972].)
- Campbell, Lyle. 1999. *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.
- Hirt, Hermann. 1927–37. *Indogermanische Grammatik*. Heidelberg: Winter.
- Meier-Brügger, Michael. 2000. *Indogermanische Sprachwissenschaft*. 8th edition. Berlin: de Gruyter. (Translated as *Indo-European Linguistics* [Berlin: de Gruyter, 2003].)
- Meillet, Antoine. 1925. *La Méthode comparative en linguistique historique*. Paris: Champion. (Translated as *The Comparative Method in Historical Linguistics* [Paris: Champion, 1967].)
- Meillet, Antoine. 1937. *Introduction à l'étude comparative des langues indo-européennes*. 8th edition. Paris: Hachette.
- Pedersen, Holger. 1959. *The Discovery of Language: Linguistic Science in the 19th Century*. Bloomington: Indiana University Press.
- Ramat, Anna Giacalone, and Paolo Ramat, eds. 1998. *The Indo-European Languages*. London: Routledge.
- Robins, R. H. 1997. *A Short History of Linguistics*. 4th edition. New York: Longman.
- Salmons, Joseph C., and Brian D. Joseph, eds. 1998. *Nostratic: Sifting the Evidence*. Amsterdam: Benjamins.
- Sihler, Andrew. 2000. *Language History: An Introduction*. Philadelphia: Benjamins.
- Szemerényi, Oswald. 1996. *Introduction to Indo-European Linguistics*. 4th edition. Oxford: Oxford University Press.

Chapter 2: Culture and Archaeology

- *Anthony, David. 1991. "The Archaeology of Indo-European Origins." *Journal of Indo-European Studies* 19:193–222.

- Benveniste, Émile. 1969. *Le Vocabulaire des institutions indo-européennes*. Paris: Minuit. (Translated as *Indo-European Language and Society* [Coral Gables: University of Miami Press, 1973].)
- Buck, Carl Darling. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: University of Chicago Press.
- Campanile, Enrico. 1977. *Ricerche di cultura poetica indoeuropea*. Pisa: Giardini.
- *Campanile, Enrico. 1990. *La ricostruzione della cultura indoeuropea*. Pisa: Giardini.
- *Cardona, George, Henry M. Hoenigswald, and Alfred Senn, eds. 1970. *Indo-European and Indo-Europeans: Papers Presented at the Third Indo-European Conference at the University of Pennsylvania*. Philadelphia: University of Pennsylvania Press.
- Dumézil, Georges. 1952. *Les Dieux des indo-européens*. Paris: Presses des Universitaires de France.
- Dumézil, Georges. 1958. *L'Ideologie tripartite des indo-européennes*. Brussels: Latomus.
- Gamkrelidze, Tamaz V., and Vyacheslav V. Ivanov. 1984. *Indoevropskij jazyk i indoevropceji*. Tbilisi: Izdatel'stvo Tbilisskogo Universiteta. (Translated as *Indo-European and the Indo-Europeans* [Berlin: de Gruyter, 1995].)
- Gimbutas, Marija. 1997. *The Kurgan Culture and the Indo-Europeanization of Europe: Selected Articles from 1952 to 1993*. Washington: Institute for the Study of Man.
- Mallory, J. P. 1989. *In Search of the Indo-Europeans: Language, Archaeology and Myth*. London: Thames and Hudson.
- Mallory, J. P., and Douglas Q. Adams, eds. 1997. *The Encyclopaedia of Indo-European Culture*. London: Fitzroy Dearborn.
- Mann, Stuart E. 1984–7. *An Indo-European Comparative Dictionary*. Hamburg: Buske.
- *Meid, Wolfgang. 1998. *Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen Gesellschaft, Innsbruck, 22.–28. September 1996*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Pokorny, Julius. 1959–69. *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.
- Puhvel, Jaan. 1987. *Comparative Mythology*. Baltimore: Johns Hopkins University Press.
- Renfrew, Colin. 1987. *Archaeology and Language: The Puzzle of Indo-European Origins*. London: Cape.
- Schmitt, Rüdiger. 1967. *Dichtung und Dichtersprache in indogermanischer Zeit*. Wiesbaden: Harrassowitz.
- *Sergent, Bernard. 1995. *Les Indo-européennes: Histoire, langues, mythes*. Paris: Payot & Rivages.
- *Watkins, Calvert. 1963. "Indo-European metrics and Archaic Irish Verse." *Celtica* 6:194–249.
- Watkins, Calvert. 1994. *Selected Writings*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Watkins, Calvert. 1995. *How to Kill a Dragon: Aspects of Indo-European Poetics*. Oxford: Oxford University Press.
- Watkins, Calvert. 2000. *The American Heritage Dictionary of Indo-European Roots*. 2nd edition. Boston: Houghton Mifflin.

Chapter 3: Phonology

- Bammesberger, Alfred, ed. 1988. *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter.
- *Beekes, Robert S. P. 1992. *Rekonstruktion und relative Chronologie. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft, Leiden, 31. August–4. September 1987*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- *Collinge, N. E. 1985. *The Laws of Indo-European*. Amsterdam: Benjamins.
- *Hopper, Paul J. 1973. "Glottalized and murmured occlusives in Indo-European." *Glotta* 7:141–66.

- *Kuryłowicz, Jerzy. 1958. *L'Accentuation des langues indo-européennes*. Wrocław: Polska Akademia Nauk.
- Lindeman, Fredrik Otto. 1987. *Introduction to the Laryngeal Theory*. Oslo: Norwegian University Press.
- Mayrhofer, Manfred. 1986. *Indogermanische Grammatik*, Band I: *Lautehre*. Heidelberg: Winter.
- *Mayrhofer, Manfred, Martin Peters, and Oskar E. Pfeiffer, eds. 1980. *Lautgeschichte und Etymologie. Akten der VI. Fachtagung der Indogermanischen Gesellschaft*, Wien, 24.–29. September 1978. Wiesbaden: Reichert.
- *Schindler, Jochem. 1977a. "A thorny problem." *Die Sprache* 23:25–35.
- *Schindler, Jochem. 1977b. "Notizen zum Sieversschen Gesetz." *Die Sprache* 23:56–65.
- Szemerényi, Oswald. 1985. "Recent developments in Indo-European linguistics." *Transactions of the Philological Society* 1985:1–71.
- Winter, Werner, ed. 1965. *Evidence for Laryngeals*. The Hague: Mouton.

Chapter 4: Morphology

- Anttila, Raimo. 1969. *Proto-Indo-European Schwabeablaut*. Berkeley: University of California Press.
- Benveniste, Émile. 1935. *Origines de la formation des noms en indo-européen*. Paris: Maisonneuve.
- Kuryłowicz, Jerzy. 1927. "o indo-européen et h hittite." In *Symbolae Grammaticae in Honorem Ioannis Rozwadowski*. Kraków: Gebethner and Wolff, pp. 1.95–104.
- Kuryłowicz, Jerzy. 1956. *L'Apophonie en indo-européen*. Wrocław: Polska Akademia Nauk.
- *Rix, Helmut, ed. 1975. *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.–14. September 1973*. Wiesbaden: Reichert.
- Saussure, Ferdinand de. 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipzig: Teubner.
- *Schlerath, Bernfried. 1985. *Grammatische Kategorien: Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft, Berlin, 20.–25. Februar 1983*. Wiesbaden: Reichert.

Chapter 5: Verb

- Jasanoff, Jay. 1978. *Stative and Middle in Indo-European*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Jasanoff, Jay. 2003. *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.
- *Narten, Johanna. 1964. *Die sigmatischen Aoriste im Veda*. Wiesbaden: Harrassowitz.
- *Narten, Johanna. 1968. "Zum 'proterodynamischen' Wurzelpräsens." In J. C. Heesterman et al. (eds.), *Pratidānam: Indian, Iranian and Indo-European Studies Presented to F. B. J. Kuiper on his Sixtieth Birthday*. The Hague: Mouton, pp. 9–19.
- Rix, Helmut, ed. 2001. *LIV: Lexikon der indogermanischen Verben*. 2nd edition. Wiesbaden: Reichert.
- Watkins, Calvert. 1969. *Indogermanische Grammatik*, Band III, Teil 1: *Geschichte der indogermanischen Verbalflexion*. Heidelberg: Winter.
- *Watkins, Calvert. 1971. "Hittite and Indo-European Studies: The denominative statives in *-ē-." *Transactions of the Philological Society* 1971:51–93.

Chapter 6: Noun

- * Benveniste, Émile. 1948. *Noms d'agent et noms d'action en indo-européen*. Paris: Maisonneuve.
- * Darms, George. 1978. *Schwäher und Schwager, Hahn und Hubn: Die Vṛddhi-Ableitungen im Germanischen*. Munich: Kitzinger.
- * Eichner, Heiner. 1973. "Die Etymologie von heth. *mehur*." *Münchener Studien zur Sprachwissenschaft* 31:53–107.
- * Hoffmann, Karl. 1955. "Ein grundsprachliches Possessivsuffix." *Münchener Studien zur Sprachwissenschaft* 6:35–40.
- Nussbaum, Alan. 1986. *Head and Horn in Indo-European*. Berlin: de Gruyter.
- Rieken, Elisabeth. 1999. *Untersuchungen zur nominalen Stammbildung des Hethitischen*. Wiesbaden: Harrassowitz.
- * Schindler, Jochem. 1972. "L'Apophonie des noms-racines." *Bulletin de la Société de Linguistique de Paris* 67:31–8.
- * Schindler, Jochem. 1975a. "L'Apophonie des thèmes indo-européennes en *-r/n-*." *Bulletin de la Société de Linguistique de Paris* 70:1–10.
- * Schindler, Jochem. 1975b. "Zum Ablaut der neutralen *s*-Stämme des Indogermanischen." In H. Rix (ed.), *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft*. Wiesbaden: Reichert, pp. 259–67.

Chapter 7: Pronouns and Other Parts of Speech

- Cowgill, Warren. 1960. "Greek *ou* and Armenian *oč*." *Language* 36:347–50.
- Gvozdanović, Jadranka, ed. 1992. *Indo-European Numerals*. Berlin: de Gruyter.
- Katz, Joshua T. 2003. *Studies in Indo-European Personal Pronouns*. Göttingen: Vandenhoeck und Ruprecht.
- Schmidt, Gernot. 1978. *Stammbildung und Flexion der indogermanischen Personalpronomina*. Wiesbaden: Harrassowitz.
- Szemerényi, Oswald. 1960. *Studies in the Indo-European System of Numerals*. Heidelberg: Winter.

Chapter 8: Syntax

- Delbrück, Berthold. 1893–1900. *Vergleichende Syntax der indogermanischen Sprachen*. Strassburg: Trübner.
- Garrett, Andrew. 1990. *The Syntax of Anatolian Pronominal Clitics*. PhD dissertation, Harvard University.
- Hale, Mark. 1987. "Wackernagel's Law and the language of the Rigveda." In C. Watkins (ed.), *Studies in Memory of Warren Cowgill (1929–1985)*. Berlin: de Gruyter, pp. 38–50.
- Inkelas, Sharon, and Draga Zec, eds. 1990. *The Phonology–Syntax Connection*. Chicago: University of Chicago Press.
- Wackernagel, Jacob. 1892. "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1:333–436.
- Wackernagel, Jacob. 1920. *Vorlesungen über Syntax*. Basel: Birkhäuser.
- Watkins, Calvert. 1963. "Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb." *Celtica* 6:1–49.
- Watkins, Calvert. 1976. "Towards Indo-European syntax: Problems and pseudo-problems." In S. Steever et al. (eds.), *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistics Society, pp. 305–26.

Chapter 9: Anatolian

- Gambel, Halet. 1999. *Corpus of Hieroglyphic Luwian Inscriptions*, vol. II: *Karatepe-Aslantaş*. Berlin: de Gruyter.
- Carruba, Onofrio. 1970. *Das Palaische: Texte, Grammatik, Lexikon*. Wiesbaden: Harrassowitz.
- Friedrich, Johannes. 1974. *Hethitisches Elementarbuch*. 2nd edition. Heidelberg: Winter.
- *Garrett, Andrew. 1992. "Topics in Lycian syntax." *Historische Sprachforschung* 105:200–12.
- Gusmani, Roberto. 1964. *Lydisches Wörterbuch*. Heidelberg: Winter. (Supplements: *Ergänzungsband, Lieferung 1*, 1980; *Ergänzungsband, Lieferung 2*, 1982; *Ergänzungsband, Lieferung 3*, 1986.)
- Hawkins, John David. 2000. *Corpus of Hieroglyphic Luwian Inscriptions*, vol. I: *Inscriptions of the Iron Age*. Berlin: de Gruyter.
- *Kimball, Sara E. 1999. *Hittite Historical Phonology*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Kronasser, Heinz. 1956. *Vergleichende Laut- und Formenlehre des Hethitischen*. Heidelberg: Winter.
- Laroche, Emmanuel. 1971. *Catalogue des textes hittites*. Paris: Klincksieck.
- *Melchert, H. Craig. 1987. "PIE velars in Luvian." In C. Watkins (ed.), *Studies in Memory of Warren Cowgill (1929–1985)*. Berlin: de Gruyter, pp. 182–204.
- Melchert, H. Craig. 1993. *Cuneiform Luvian Lexicon*. Chapel Hill: self-published.
- Melchert, H. Craig. 1994. *Anatolian Historical Phonology*. Amsterdam: Rodopi.
- Melchert, H. Craig. 2004. *A Dictionary of the Lycian Language*. Ann Arbor: Beech Stave.
- Neumann, Günter. 1969. "Lykisch." In B. Spuler (ed.), *Handbuch der Orientalistik: Altkeinasische Sprachen*. Leiden: Brill, pp. 358–96.
- *Otten, Heinrich, and Vladimír Souček. 1969. *Ein althethitisches Ritual für das Königspaar*. Wiesbaden: Harrassowitz.
- Puhvel, Jaan. 1984–. *Hittite Etymological Dictionary*. Berlin: de Gruyter.
- Starke, Frank. 1984. *Die keilschrift-luwischen Texte in Umschrift*. Wiesbaden: Harrassowitz.
- Starke, Frank. 1990. *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*. Wiesbaden: Harrassowitz.
- Tischler, Johann. 1977–. *Hethitisches etymologisches Glossar*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- *Yoshida, Kazuhiko. 1990. *The Hittite Mediopassive Endings in -ri*. Berlin: de Gruyter.

Chapter 10: Indic

- Aufrecht, Theodor. 1877. *Die Hymnen des Rigveda*. 2nd edition. Bonn: Marcus.
- Delbrück, Berthold. 1886. *Altindische Syntax*. Halle: Verlag der Waisen Hauses.
- Geldner, Karl Friedrich. 1951–7. *Der Rig-Veda*. Cambridge, MA: Harvard University Press.
- Grassmann, Hermann. 1873. *Wörterbuch zum Rig-Veda*. Leipzig: Brockhaus. (6th edition with revisions by M. Kozianka [Wiesbaden: Harrassowitz], 1996.)
- *Hale, Mark. 1990. "Preliminaries to the study of the relationship between sandhi and syntax in the language of the Rigveda." *Münchener Studien zur Sprachwissenschaft* 51:77–96.
- *Hale, Mark. 1995. *Wackernagel's Law in the Language of the Rigveda*. Manuscript, Concordia University, Montréal.
- Hoffmann, Karl. 1975–92. *Aufsätze zur Indo-Iranistik*. Wiesbaden: Reichert.
- *Jamison, Stephanie. 1991. *The Ravenous Hyenas and the Wounded Sun: Myth and Ritual in Ancient India*. Ithaca: Cornell University Press.
- Macdonell, Arthur Anthony. 1910. *A Vedic Grammar*. Strassburg: Trübner.
- Macdonell, Arthur Anthony. 1916. *A Vedic Grammar for Students*. Oxford: Clarendon.

- Macdonell, Arthur Anthony. 1917. *A Vedic Reader for Students*. Oxford: Clarendon.
- *Mayrhofer, Manfred. 1966. *Die Indo-Arier im Alten Vorderasien*. Wiesbaden: Harrassowitz.
- Mayrhofer, Manfred. 1986–2001. *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg: Winter.
- Oberlies, Thomas. 2001. *Pāli: A Grammar of the Language of the Theravāda Tipiṭaka*. Berlin: de Gruyter.
- Renou, Louis. 1955–69. *Études védiques et pāṇinéennes*. Paris: Boccard.
- Wackernagel, Jacob, and Albert Debrunner. 1896–. *Altindische Grammatik*. Göttingen: Vandenhoeck und Ruprecht.
- Whitney, William Dwight. 1896. *Sanskrit Grammar*. Leipzig: Breitkopf und Härtel.

Chapter 11: Iranian

- Bartholomae, Christian. 1904. *Altiranisches Wörterbuch*. Strassburg: Trübner. (Reprint with the 1906 supplement [Berlin: de Gruyter], 1979.)
- Geldner, Karl F. 1886–96. *Avesta: The Sacred Books of the Parsis*. Stuttgart: Kohlhammer.
- Hoffmann, Karl, and Bernhard Forssman. 1996. *Avestische Laut- und Formenlehre*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Humbach, Helmut. 1991. *The Gāthās of Zarathuštra and the Other Old Avestan Texts*. Heidelberg: Winter.
- Insler, Stanley. 1975. *The Gāthās of Zarathustra*. Leiden: Brill.
- Kellens, Jean. 1974. *Les Noms-racines de l'Avesta*. Wiesbaden: Reichert.
- Kellens, Jean. 1984. *Le Verbe avestique*. Wiesbaden: Reichert. (Supplement: *Liste du verbe avestique*, 1994.)
- Kent, Roland. 1953. *Old Persian: Grammar, Texts, Lexicon*. 2nd edition. New Haven: American Oriental Society.
- Reichelt, Hans. 1909. *Awestisches Elementarbuch*. Heidelberg: Winter.
- Schmitt, Rudiger, ed. 1989. *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert.
- Yarshater, Ehsan, ed. 1982–. *Encyclopedia Iranica*. London: Eisenbrauns.

Chapter 12: Greek

- Allen, W. Sidney. 1987. *Vox Graeca: A Guide to the Pronunciation of Ancient Greek*. 3rd edition. Cambridge, UK: Cambridge University Press.
- *Bartonek, Antonín. 2003. *Handbuch des mykenischen Griechisch*. Heidelberg: Winter.
- Buck, Carl Darling. 1968. *The Greek Dialects: Grammar, Selected Inscriptions, Glossary*. Chicago: University of Chicago Press.
- Chantraine, Pierre. 1953–8. *Grammaire homérique*. Paris: Klincksieck.
- Chantraine, Pierre. 1968–80. *Dictionnaire étymologique de la langue grecque*. Paris: Klincksieck.
- Devine, A. M., and Laurence D. Stephens. 1994. *The Prosody of Greek Speech*. Oxford: Oxford University Press.
- Duhoux, Yves. 1984. *Introduction aux dialectes grecs anciens*. Louvain-la-Neuve: Cabay.
- Frisk, Hjalmar. 1960–72. *Griechisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Heubeck, Alfred, et al., eds. 1988–92. *A Commentary on Homer's Odyssey*. Oxford: Oxford University Press.
- *Hooker, J. T. 1980. *Linear B: An Introduction*. Bristol: Bristol Classical Press.
- *Horrocks, Geoffrey. 1997. *Greek: A History of the Language and its Speakers*. London: Longman.
- Kirk, G. S., ed. 1985–93. *The Iliad: A Commentary*. Cambridge, UK: Cambridge University Press.

- Lejeune, Michel. 1972. *Phonétique historique du mycénien et du grec ancien*. Paris: Klincksieck.
- *Meier-Brügger, Michael. 1992. *Griechische Sprachwissenschaft*. Berlin: de Gruyter.
- Meillet, Antoine, and J. Vendryes. 1968. *Traité de grammaire comparée des langues classiques*. 4th edition. Paris: Champion.
- Peters, Martin. 1980. *Untersuchungen zur Vertretung der indogermanischen Laryngale im Griechischen*. Vienna: Österreichische Akademie der Wissenschaften.
- Risch, Ernst. 1974. *Wortbildung der homerischen Sprache*. 2nd edition. Berlin: de Gruyter.
- Rix, Helmut. 1992. *Historische Grammatik des Griechischen: Laut- und Formenlehre*. 2nd edition. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Schmitt, Rüdiger. 1977. *Einführung in die griechischen Dialekte*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Schwyzler, Eduard. 1939–71. *Griechische Grammatik*. Munich: Beck.
- Sihler, Andrew. 1995. *New Comparative Grammar of Greek and Latin*. New York: Oxford University Press.
- Ventris, Michael, and James Chadwick. 1973. *Documents in Mycenaean Greek*. Cambridge, UK: Cambridge University Press.

Chapter 13: Italic

- *Allen, W. Sidney. 1989. *Vox Latina: A Guide to the Pronunciation of Classical Latin*. Cambridge, UK: Cambridge University Press.
- Baldi, Philip. 1999. *The Foundations of Latin*. Berlin: de Gruyter.
- Buck, Carl Darling. 1928. *A Grammar of Oscan and Umbrian*. Boston: Ginn.
- Ernout, Alfred. 1947. *Recueil de textes latins archaïques*. 2nd edition. Paris: Klincksieck.
- Ernout, Alfred, and Antoine Meillet. 1979. *Dictionnaire étymologique de la langue latine*. 4th edition revised by Jacques André. Paris: Klincksieck.
- Giacomelli, Gabriella. 1963. *La lingua falisca*. Florence: Olshcki.
- Leumann, Manu. 1977. *Lateinische Laut- und Formenlehre*. Munich: Beck.
- Marinetti, Anna. 1985. *Le iscrizioni sudpicene*. Florence: Olshcki.
- Meiser, Gerhard. 1986. *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Meiser, Gerhard. 1998. *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- *Meyer-Lübke, Wilhelm. 1935. *Romanisches etymologisches Wörterbuch*. 3rd edition. Heidelberg: Winter.
- Palmer, L. R. 1954. *The Latin Language*. London: Faber & Faber.
- Poccetti, Paolo. 1979. *Nuovi documenti italici a complemento del Manuale di E. Vetter*. Pisa: Giardini.
- Poultney, James. 1959. *The Bronze Tables of Iguvium*. Baltimore: American Philological Association.
- Rix, Helmut. 2002. *Sabellische Texte: Die Texte des Oskischen, Umbrischen, und Südpikenischen*. Heidelberg: Winter.
- Schrijver, Peter. 1991. *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam: Rodopi.
- *Sommer, Ferdinand. 1977. *Handbuch der lateinischen Laut- und Formenlehre*, vol. 1, *Lautlehre*. 4th edition, revised by Raimund Pfister. Heidelberg: Winter.
- Untermann, Jürgen. 1999. *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- *Väänänen, Veikko. 1981. *Introduction au latin vulgaire*. 3rd edition. Paris: Klincksieck.
- Vetter, Emil. 1953. *Handbuch der italischen Dialekte*. Heidelberg: Winter.
- Vine, Brent. 1993. *Studies in Archaic Latin Inscriptions*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Wachter, Rudolf. 1987. *Altlateinische Inschriften*. Bern: Lang.

Chapter 14: Celtic

- Duval, Paul-Marie, ed. 1985–2002. *Recueil des inscriptions gauloises*. Paris: Centre National de la Recherche Scientifique.
- Evans, D. Simon. 1964. *A Grammar of Middle Welsh*. Dublin: Institute for Advanced Studies.
- Forsyth, Katherine. 1997. *Language in Pictland*. Utrecht: Keltische Draak.
- Jackson, Kenneth. 1953. *Language and History in Early Britain: A Chronological Survey of the Brittonic Languages, First to Twelfth Century A.D.* Cambridge, MA: Harvard University Press.
- Lewis, Henry, and Holger Pedersen. 1937. *A Concise Comparative Celtic Grammar*. Göttingen: Vandenhoeck und Ruprecht.
- *McCone, Kim. 1991. *The Indo-European Origins of the Old Irish Nasal Presents, Subjunctives and Futures*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- McCone, Kim. 1996. *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change*. Maynooth: Department of Old Irish, St. Patrick's College.
- McManus, Damian. 1991. *A Guide to Ogam*. Maynooth: An Sagart.
- Meid, Wolfgang. 1992. *Gaulish Inscriptions*. Budapest: Archaeological Institute of the Hungarian Academy of Sciences.
- Meid, Wolfgang. 1994. *Celtiberian Inscriptions*. Budapest: Archaeological Institute of the Hungarian Academy of Sciences.
- Pedersen, Holger. 1909–13. *Vergleichende Grammatik der keltischen Sprachen*. Göttingen: Vandenhoeck und Ruprecht.
- *Schmidt, Karl Horst, ed. 1977. *Indogermanisch und Keltisch. Kolloquium der Indogermanischen Gesellschaft am 16. und 17. Februar in Bonn*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Schrijver, Peter. 1995. *Studies in British Celtic Historical Phonology*. Amsterdam: Rodopi.
- *Strachan, John. 1949. *Old-Irish Paradigms and Selections from the Old-Irish Glosses*. Dublin: Royal Irish Academy.
- Thurneysen, Rudolf. 1980. *A Grammar of Old Irish*. 2nd edition, translated and revised by D. A. Binchy and Osborn Bergin. Dublin: Institute for Advanced Studies.
- Untermann, Jürgen. 1997. *Monumenta Linguarum Hispanicarum*. Vol. 4: *Die tartessischen, keltiberischen, und lusitanischen Inschriften*. Vol. 5: *Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Reichert.
- Vendryes, Jean, and P.-Y. Lambert. 1959–. *Lexique étymologique de l'irlandais ancien*. Dublin: Institute for Advanced Studies.
- *Villar, F. 1995. *A New Interpretation of Celtiberian Grammar*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- *Watkins, Calvert. 1962. *Indo-European Origins of the Celtic Verb I. The Sigmatic Aorist*. Dublin: Institute for Advanced Studies.

Chapter 15: Germanic

- Antonsen, Elmer H. 1975. *A Concise Grammar of the Older Runic Inscriptions*. Tübingen: Niemeyer.
- Bammesberger, Alfred. 1986. *Der Aufbau des urgermanischen Verbalsystems*. Heidelberg: Winter.
- Bammesberger, Alfred. 1990. *Die Morphologie des urgermanischen Nomens*. Heidelberg: Winter.
- Braune, Wilhelm. 1981. *Gotische Grammatik*. 19th edition, revised by Ernst A. Ebbinghaus. Tübingen: Niemeyer.
- Braune, Wilhelm. 1987. *Althochdeutsche Grammatik*. 14th edition, revised by H. Eggers. Tübingen: Niemeyer.

- Braune, Wilhelm. 1994. *Althochdeutsches Lesebuch*. 17th edition, revised by Ernst A. Ebbinghaus. Tübingen: Niemeyer.
- Campbell, Alistair. 1959. *Old English Grammar*. Oxford: Oxford University Press.
- *Cathey, James E. 2002. *Heliand: Text and Commentary*. Morgantown: West Virginia University Press.
- de Vries, Jan. 1962. *Altnordisches etymologisches Wörterbuch*. 2nd edition. Leiden: Brill.
- Gordon, E. V. 1962. *An Introduction to Old Norse*. 2nd edition. Oxford: Oxford University Press.
- Green, Dennis H. 1998. *Language and History in the Early Germanic World*. Cambridge: Cambridge University Press.
- Holthausen, F. 1921. *Altsächsisches Elementarbuch*. Heidelberg: Winter.
- Holthausen, F. 1934. *Altenglisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Kluge, Friedrich. 2002. *Etymologisches Wörterbuch der deutschen Sprache*. 24th edition, revised by Elmar Seebold. Berlin: de Gruyter.
- Krause, Wolfgang. 1968. *Handbuch des Gotischen*. 3rd edition. Munich: Beck.
- Krause, Wolfgang. 1971. *Die Sprache der urnordischen Runeninschriften*. Heidelberg: Winter.
- Krause, Wolfgang, and Herbert Jankuhn. 1966. *Die Runeninschriften im älteren Futhark*. Göttingen: Vandenhoeck und Ruprecht.
- *Kyes, Robert L. 1969. *The Old Low Franconian Psalms and Glosses*. Ann Arbor: University of Michigan Press.
- Lehmann, Winfred P. 1986. *A Gothic Etymological Dictionary*. Leiden: Brill.
- Nielsen, Hans Frede. 1989. *The Germanic Languages: Origins and Early Dialectal Interrelations*. Tuscaloosa: University of Alabama Press.
- Nielsen, Hans Frede. 2000. *The Early Runic Language of Scandinavia*. Heidelberg: Winter.
- Noreen, Adolf. 1923. *Altisländische und altnorwegische Grammatik*. 4th edition. Halle: Niemeyer.
- Prokosch, Edward. 1939. *A Comparative Germanic Grammar*. Philadelphia: Linguistic Society of America.
- Robinson, Orrin W. III. 1992. *Old English and Its Closest Relatives: A Survey of the Germanic Languages*. Stanford: Stanford University Press.
- Seebold, Elmar. 1970. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague: Mouton.
- Streitberg, Wilhelm. 2000. *Die gotische Bibel*. 7th edition. Heidelberg: Winter.

Chapter 16: Armenian

- Clackson, James. 1994. *The Linguistic Relationship between Armenian and Greek*. Oxford: Blackwell.
- Godel, Robert. 1975. *An Introduction to the Study of Classical Armenian*. Wiesbaden: Reichert.
- Hübschmann, Heinrich. 1895–7. *Armenische Grammatik, Teil 1: Armenische Etymologie*. Leipzig: Breitkopf und Härtel.
- Klingenschmitt, Gert. 1982. *Das altarmenische Verbum*. Wiesbaden: Reichert.
- Lamberterie, Charles. 1992. *Introduction à l'arménien classique*. LALIES 10:234–89.
- Meillet, Antoine. 1936. *Esquisse d'une grammaire comparée de l'arménien classique*. 2nd edition. Vienna: Mékhitaristes.
- Meillet, Antoine. 1962–77. *Études de linguistique et de philologie arméniennes*. Louvain: Imprimerie Orientaliste.
- Olsen, Birgit Anette. 1999. *The Noun in Biblical Armenian: Origin and Word-Formation*. Berlin: de Gruyter.
- Schmitt, Rüdiger. 1981. *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

- *Solta, Georg R. 1960. *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*. Vienna: Mechitharisten-Buchdruck.
- *Thomson, Robert. 1989. *An Introduction to Classical Armenian*. 2nd edition. Delmar: Caravan.

Chapter 17: Tocharian

- Adams, Douglas Q. 1988. *Tocharian Historical Phonology and Morphology*. New Haven: American Oriental Society.
- Adams, Douglas Q. 1999. *A Dictionary of Tocharian B*. Amsterdam: Rodopi.
- Krause, Wolfgang, and Werner Thomas. 1960–4. *Tocharisches Elementarbuch*. Heidelberg: Winter.
- Mallory, J. P., and Victor H. Mair. 2000. *The Tarim Mummies: Ancient China and the Mystery of the Earliest Peoples from the West*. New York: Thames & Hudson.
- Pinault, Georges-Jean. 1989. "Introduction au tokharien." *LALIES* 7:1–224.
- Ringe, Donald. 1996. *On the Chronology of Sound Changes in Tocharian*, vol. 1: *From Proto-Indo-European to Proto-Tocharian*. New Haven: American Oriental Society.
- *Schlerath, Bernfried, ed. 1994. *Tocharisch. Akten der Fachtagung der Indogermanischen Gesellschaft, Berlin, September 1990*. Reykjavik: Málvísindastofnun Háskóla Íslands.
- Sieg, Emil, and Wilhelm Siegling. 1921. *Tocharische Sprachreste*. Berlin: de Gruyter.
- Sieg, Emil, and Wilhelm Siegling. 1949–53. *Tocharische Sprachreste, Sprache B*. Göttingen: Vandenhoeck und Ruprecht.
- Thomas, Werner. 1983. *Tocharische Sprachreste, Sprache B*. Göttingen: Vandenhoeck und Ruprecht.
- van Windekens, A. J. 1976–82. *Le tokharien confronté avec les autres langues indo-européennes*. Louvain: Centre international de dialectologie générale.

Chapter 18: Balto-Slavic

- Carlton, Terence R. 1990. *Introduction to the Phonological History of the Slavic Languages*. Columbus: Slavica.
- Endzelins, Jānis. 1971. *Jānis Endzelins' Comparative Phonology and Morphology of the Baltic Languages*. Translated by William R. Schmalstieg and Benjamin Jēgers. The Hague: Mouton.
- Fraenkel, Ernst. 1962–5. *Litauisches etymologisches Wörterbuch*. Heidelberg: Winter.
- *Illič-Svityč, V. M. 1979. *Nominal Accentuation in Baltic and Slavic*. Translated by R. L. Leed and R. F. Feldstein. Cambridge, MA: MIT Press.
- Leskien, August. 1990. *Handbuch der altpolnischen (altkirchenslavischen) Sprache*. 10th edition, revised by Johannes Schröpper. Heidelberg: Winter.
- Lunt, Horace. 2001. *Old Church Slavonic Grammar*. 7th edition. Berlin: de Gruyter.
- Maziulis, Vytautas. 1988–97. *Praus kalbos etimologijos žodynas*. Vilnius: Mokslas.
- Meillier, Antoine. 1965. *Le slave commun*. 2nd edition. Paris: Champion.
- *Poljakov, Oleg. 1995. *Das Problem der balto-slavischen Sprachgemeinschaft*. Frankfurt: Lang.
- Schmalstieg, William R. 1974. *An Old Prussian Grammar*. University Park: Pennsylvania State University Press.
- Senn, Alfred. 1957–66. *Handbuch der litauischen Sprache*. Heidelberg: Winter.
- Stang, Christian. 1942. *Das slavische und baltische Verbum*. Oslo: Dybwad.
- Stang, Christian. 1965. *Slavonic Accentuation*. 2nd edition. Oslo: Universitetsforlaget.
- Stang, Christian. 1966. *Vergleichende Grammatik der baltischen Sprachen*. Oslo: Universitetsforlaget. (Supplement: *Ergänzungsband*, 1975.)
- Toporov, V. N. 1975–. *Prusskij jazyk*. Moscow: Nauka.

- Trautmann, R. 1910. *Die altpreussischen Sprachdenkmäler*. Göttingen: Vandenhoeck und Ruprecht.
- Trubachev, O. N. 1974–. *Ėtimologičeskij slovar' slavjanskix jazykov*. Moscow: Nauka.
- Vaillant, André. 1950–77. *Grammaire comparée des langues slaves*. Lyon: IAC.
- Vasmer, Max. 1953–8. *Russisches etymologisches Wörterbuch*. Heidelberg: Winter. (Translated into Russian and revised as *Ėtimologičeskij slovar' russkogo jazyka* [Moscow: Progress], 1964.)
- *Winter, Werner. 1976. "The distribution of short and long vowels in stems of the type Lith. *estī* : *vēstī* : *mēstī* and OCS *jaŝti* : *veŝti* : *mēstī* in Baltic and Slavic languages." In Jacek Fisiak (ed.), *Recent Developments in Historical Phonology*. The Hague: Mouton, pp. 431–46.
- Zinkevičius, Zigmas. 1996. *The History of the Lithuanian Language*. Vilnius: Mokslo ir enciklopedijų leidykla.

Chapter 19: Albanian

- Camaj, Martin. 1984. *Albanian Grammar with Exercises, Chrestomathy and Glossaries*. Wiesbaden: Harrassowitz.
- Demiraj, Bardhyl. 1997. *Albanische Etymologien: Untersuchungen zum albanischen Wortschatz*. Amsterdam: Rodopi.
- Demiraj, Shaban. 1993. *Historische Grammatik der albanischen Sprache*. Vienna: Österreichische Akademie der Wissenschaften.
- Fox, Leonard. 1989. *Kanuni i Lekë Dukagjinit: The Code of Lekë Dukagjini*. New York: Gjonlekaj.
- Hamp, Eric P. 1977. "A far-out equation." In Irén Hegedüs et al., eds., *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*. Washington: Institute for the Study of Man, pp. 94–105.
- Huld, Martin E. 1984. *Basic Albanian Etymologies*. Columbus: Slavica.
- Meyer, Gustav. 1891. *Etymologisches Wörterbuch der albanesischen Sprache*. Strassburg: Trübner.
- Orel, Vladimir. 1998. *Albanian Etymological Dictionary*. Leiden: Brill.
- Orel, Vladimir. 2000. *A Concise Historical Grammar of the Albanian Language*. Leiden: Brill.
- Sanz Ledesma, Manuel. 1996. *El albanés: Gramática, historia, textos*. Madrid: Ediciones Clásicas.

Chapter 20: Fragmentary Languages

- Brixhe, Claude, and Michel Lejeune. 1984. *Corpus des inscriptions paléo-phrygiennes*. Paris: Recherche sur les civilisations.
- Detschew, Dimiter. 1957. *Die thrakischen Sprachreste*. Vienna: Rohrer.
- Haas, Otto. 1966. *Die phrygischen Sprachdenkmäler*. Sofia: Académie bulgare des sciences.
- Kallérís, J. N. 1954–76. *Les anciens Macédoniens: Étude linguistique et historique*. Athens: Institut français d'Athènes.
- Katičić, Radoslav. 1976. *Ancient Languages of the Balkans*. The Hague: Mouton.
- Krahe, Hans. 1955–64. *Die Sprache der Illyrier*. Wiesbaden: Harrassowitz.
- Lejeune, Michel. 1974. *Manuel de la langue vénète*. Heidelberg: Winter.
- Neumann, Günter. 1988. *Phrygisch und Griechisch*. Vienna: Österreichische Akademie der Wissenschaften.
- Orel, Vladimir. 1997. *The Language of the Phrygians*. Delmar: Caravan.
- Untermann, Jürgen. 1997. *Monumenta Linguarum Hispanicarum*, vol. 4: *Die tartessischen, keltiberischen, und lusitanischen Inschriften*. Wiesbaden: Reichert.

Word Index

This index contains reconstructed and attested forms for which historical information or paradigms are given. The order of branches and languages is the same as the chapter order. Wherever possible, inflected forms of the same word are listed together; for PIE, inflected and derived forms are normally listed under their root. Phrases are given their own entries. Except for PIE (see below), words are listed in English alphabetical order, and diacritics and parentheses are ignored. Where they occur, β δ γ λ θ χ are alphabetized as *b d g l th x*, and φ δ ι ν η γ as *e th dh hu ng*.

I. Proto-Indo-European

Alphabetical order: *a ā b bh d dh e ē g ġ gh ĝh gʷ h H h₁ h₂ h₃ i ī j k k̃ kʷ l ʎ m/n n̄ n̅ o ō p r/r s t u ū x z*

**abel-*, 40, 72
 **agʷnos*, 72, 116
 **agro-*, 356
 **ajer-*, 322
 **al-jo-*, 130
 **anghos-*, 253
 **apo*, 134
 **at-*, 370
 **bel-*, 52
 **bhāgos*, 113
 **bhe-bhru-*, 112
 **bheh₂-*, 117; **bheh₂-nu*, 345; **bheh₂-ti*, 239
 **bheidh-*, 393; **bhidh-tā*, 393, 394
 **bhendh-*, 304
 **bher-*, 53, 89, 95, 322, 341, 345, 365, 394; **bher-oh₂*, 275, 372; **bhereti*, 284, 285; **bher-onti*, 183; **ébherom*, 228; **ebberet*, 91, 227, 343, 346; **bhērete*, 95; **bhéro-ih₁-n̄*, 96; **bhēr-e*, 95; **bher-ont-*, 97; **bboro-*, 341, 347

**bherġh-*, 279; **bbrġġh-*, 340; **bbrġġhent-*, 182; **bbrġġh-ro-*, 354
 **bberHġeh₂*, 366
 **bheudh-*, 360, 371; **bheudh-eti*, 188; **bhudh-tó-*, 188
 **bhi-n-d-*, 249; **bhid-to-*, 248; **bhid-nó-*, 98
 **bhlégh-my*, 110
 **bhlégh-mó(n)*, 110
 **bhos-o-*, 74
 **bhrātēr*, **bhrāter-*, 61, 112, 182, 304, 353, 355
 **bbreu-*, 112, 404
 **bhug-je-*, 249
 **bhuH-*, 74, 250, 267, 380; **bhu(H)-e-*, 250
 **daġ-*, 353
 **de*, 135
 **deh₃-*, 71, 77, 88, 118, 341, 349, 393, 408; **de-deh₃-*, **de-dh₃-*, 77, 88; **deh₃s*, 208; **dh₃-tó-*, 229; **dh₃-to-*, 248; **deh₃-rom*, 58; **déh₃-tór*, 112; **dh₃-tēr*, 111–12; **dh₃-tr-*, 57
 **deigh-ā*, 393
 **deik-s-*, 61
 **deu-ó-s*, 117; **deiu-ons*, 207, 386
 **dekm*, 56, 112, 131, 181, 340, 341, 360; **dekṃ-t-*,

112, 384; **dekmno-*, 132; **dekṃ-ti-*, 394
 **dem-* ‘build’, 77
 **dem-* ‘house’, 108; **dóm-s*, 108; **dem-s*, 77, 108; **dm(m)-és*, 108
 **dems potis*, 108, 207
 **dens-*, 162
 **derk-*, 227
 **djēus*, **djeu-*, 61, 64, 66, 104, 109, 253, 408; **deiu-*, 117; **djem*, 64, 109; **dieu-os*, 403
 **djēus ph₂tēr*, 406
 **din-*, 375
 **diuots*, 172
 **dīh₁gho-*, 182
 **dyġhu-*, 248, 354
 doru*, **deru*, 52, 108, 353, 355; **dēru-s*/dr-ēu-s* (gen. sg.), 108
 **drk-o-*, 394
 **d(u)h₁oh₁*, 105, 131; **dūi-*, 131, 174; **dūis*, 133
 **dyeh₂-ro-*, **duh₁-ro-*, 239, 342
 **dyghā-*, 394
 **-dhe* (adv. suff.), 107
 **dheġhōm*, **dheġhom-*, 109, 110, 165, 167, 354, 397; **dhġhem-*, 361, 402; **dhġhōm* (acc.), 109; **dhġhom-*, 59; **dhġhmēi*, 165; **dhġhm-en*, 110

- *dheg^wh-, 354, 393
 *dheh₁-, 71, 210, 279, 341, 371, 403; *dhe-dheh₁-mi-, 227; *dhe-dheh₁-(e)nt-, 97; *e-dheh₁-, 91; *dheh₁s-, 208; *dhh₁-tó-, 229; *dhh₁-k-, 403, 406; *dheh₁-m₁, 58; *dheh₁-ti-, 304; *dhh₁-m₁-, 210
 *dheh₁s-, 72, 239; *dhh₁s-, 57
 *dheigh₁-, 263; *dhoigh₁-, 340, 343
 *dheub-, 353; *dhub-ro-, 353; *dhub-u-, 365
 *dhghemō(n), 64, 110, 327, 383
 *dhghā-, 61, 342; *ghdhūs-, 230
 *dhg^wher-, 182, 197
 *dhg^whi-, 59
 *-dhi (adv. suff.), 107
 *-dhlo-, 118
 *-dhro-, 118
 *dh₁-to-, 209
 *dhugh₁tēr-, *dhugh₂ter-, 72, 112, 173, 204, 342
 *dhuh₁-mo-, 209
 *dhuenh₂-, 209
 *e-, 130 (for augmented forms, see under relevant root)
 *eg₁-, 156, 349, 380, 384; *egoh₂-, 127, 296, 358; *egh₂-om-, 127, 213
 *eg^wh-, 172
 *(e)-g^when-t og^whim, 26
 *-eh₂- (fem. suff.), 119
 *ei- (dem. pron. stem), 130
 *eks, 176
 *ekyos, 240, 354, 380
 *eme-, see *me-
 *en-, 134, 362, 394
 *endo, 162, 173
 *eno-, *ono-, 130, 344
 *en-ter, 134, 214
 *-(e)nt(i), 85
 *epi-, *opi-, *epi-, 134, 341, 347, 348, 408
 *erk^wos, 360
 *gal-, 117
 *ge, 135
 *gras-, 52
 *genh₁-, 77, 90, 112, 322, 349; *gi-ghh₁-, 77; *ghh₁-tó-, 229, 248; *ghénh₁-os, 112; *ghh₁-ti-, 229; *ghénh₁-to-, 366
 *gneh₁-, 117, 355, 368, 371; *ghh₁-tó-, 229, 276, 304; *gneh₃u-, 257; *(-)ghh₁-ti-, 229
 *gombhos, 366, 371, 393
 *gonu-, *genu-, 52, 108, 230; *gnu-, 393
 *ghh₂no-, *ghh₂-n-, 276, 380
 *ghe-, *gho-, 135, 374
 *ghebb₁-(e)l-, 354
 *gheluneh₂-, 72
 *ghor-dho-, 365, 366, 393, 394
 *ghort-, 264
 *ghostis, 303
 *ghous-, 181
 *ghrebbh-, *ghrobh-, 163
 *ghrebbh₂-, 191
 *ghans-, 365, 371
 *ghem-, 384; *gheimon-, 393, 395
 *ghel-, 207, 275; *ghel-to-, 365
 *ghl-to-, 365
 *gher-, 322
 *gherzdheh₂-, *ghy₂zdo-, 394
 *ghés-ör, 112; *ghes-r-, 168; *ghes-ro-, 168
 *ghew-, 53, 263, 354
 *ghlō-ro-, 402
 *g^wem-, 98, 253, 262, 269, 360; *g^wm₁-, 355; *g^wm₁-tó-, 98; *g^wm₁-ti-, 229
 *g^wem-, 52, 168, 275, 355, 370, 375, 386; *g^wenh₂-, 365
 *g^weneh₂-, 119
 *g^werh₃-, 116, 365, 393; *g^wor(h₃)os, 116
 *g^wih₃-, 264; *g^wih₃-yo-, 380
 *g^wieh₂-, 116; *g^wih₂-ó-, 116
 *g^wieh₃u-ont-, 361
 *g^wōus, 61, 104, 181, 190; *g^wōu-, 173, 290; *g^wōu-, 60, 341, 353; *g^wōm, 64
 *g^wrh₂u-, 111; *g^wreh₂u-ō(n), 111
 *g^wheid-je-, 302
 *g^when-, 26-7, 53, 91, 98, 341; *g^when-ti, 161, 188; *g^whi-g^wh₁-h₁se-ti, 91; *g^wh₁-tó-, 98; *g^whén-s-, 109; *g^whonos, 302
 *g^wher-, 117; *g^wher-mo-, 117, 341, 393; *g^whor-mo-, 117, 302
 *-Hon-, 111
 *h₁ed-, 162, 347; *h₁éd-si-, *h₁éd-ti, 63; *h₁edonti, 172; *h₁öd-éje-, 345, 348; *h₁d-yon-, 342, 347
 *h₁ei-, 172; *h₁ei-dhi, 167, 188, 374; *h₁ei₁tōd, 403
 *h₁es-, 71, 87 (paradigm), 95, 308, 345; *h₁es-mi-, 394, 396; *h₁es-si-, *h₁esi, 63, 345; *h₁esti, 57-8, 60, 289, 375, 398; *h₁s-énti, 304; *h₁es-m₁-, 396; *h₁es-e-, 250; *e-h₁es-, 345; *e-h₁es-ti, 195; *h₁s-ieh₁-m, 96 (paradigm), 233; *h₁s-ieh₁-s, 251, 261; *h₁s-ih₁-me-, 233; *h₁s-ih₁-te-, 251; *h₁s-ent-, 97; *en-h₁esti, 393, 398
 *h₁esu-, 123; *h₁su- (pref.), 123, 189
 *h₁eus-e-, 254; *h₁us-no-, 188
 *h₁leudh-, 92; *(é)h₁ludh-e-, 92; *h₁ludh-et, 354
 *h₁neh₃m₁ (*h₁nom₁), 111, 248; *h₁neh₃m₁-, 360; *h₁neh₃m₁-, 90
 *h₁reg^w-, 229; *h₁reg^wos, 342, 348
 *h₁reudh-, 71, 91, 370; *h₁rudh-, 370; *h₁rudh-ro-, 121, 353, 355, 356; *h₁rudh-eh₁-, 91
 *-h₂ (fem. suff.), 118, 119
 *-h₂ (1st sg.), 85, 86, 96
 *h₂eg₁-, 348, 353, 355; *h₂eg₁-, 249, 250, 345; *h₂eg₁-o-mes, 358; *h₂eg₁-e-e-, 250; *h₂eg₁-e-tōd, 403
 *h₂eg₁-s-, 380
 *h₂ek-, 393
 *h₂ek-men-, 365
 *h₂el-, 322
 *h₂em-, 342, 362
 *h₂enh₁mo-, 248
 *h₂ens-, 333
 *h₂ent-, 57-8, 71; *h₂enti, 325, 348
 *h₂erh₁-, 118, 383; *h₂erh₁-trom, 118, 342
 *h₂(e)u, 134
 *h₂eug-, 353, 356

- **h₂euh₂os*, 342
 **h₂eus₂*, 315, 383; **h₂éus-ōs*, 109, 380; **h₂eus-r*, 380
 **h₂euso-*, 370
 **h₂kous-*, 61, 74, 230, 315
 **h₂ṛbhi*, **h₂ṛbhi*, 239, 276
 **h₂melg-*, 290
 **h₂nēr*, 57, 209, 402; **h₂ner-*, 71, 189, 267, 342, 394
 **h₂oju*, 108, 111, 133, 239; **h₂oju-*, 322; **h₂nt-Hō(n)*, 111
 **h₂oi-*, 370
 **h₂ṛg-ró-*, 121
 **h₂ṛtkos*, 59, 74, 116, 276
 **h₂ster*, 302
 **h₂uk^w-s-én*, 109
 **h₂uēks-*, **h₂uōks-*, 164, 182
 **h₂uēh₁-*, 195; **h₂uēh₁-ṛt-os*, 189, 195, 209; **h₂uēh₁-nto-*, 64, 354; **h₂uēh₁-ju-*, 195
 **h₂uēs-*, 342
 **h₂bhel-*, 72, 229
 **h₂ed-*, 71, 342
 **h₂ek^w-*, 108, 315, 348, 353, 355, 361, 380; **h₂(e)k^wih₁*, 105, 393
 **h₂emh₂-*, 241
 **h₂er-*, 162; **h₂er(o)n-*, 342
 **h₂erbh-*, 57-8
 **h₂ezd-*, 342
 **h₂meigh-*, 71, 341
 **h₂neid-je-*, 345
 **h₂reḡ-*, 279, 322; **h₂rēgs*, 60
 *-i (primary active marker), 85
 *-ih₂, 64-5, 354-5
 *-iō-, see **jo-*
 *-iko-, 121, 370
 *-isko-, 121
 *-isto- (ordinal suff.), 122, 132
 *-isto- (superl. suff.), 122
 *-it- (n. suff.), 112
 **iaḡ-*, 109, 117, 228, 230
 **ieh₁-i-ih-* (opt. suff.), 96
 **ieh₂-* (fem. suff.), 120
 **iēk^w-r*, 108
 **ienh₂er-*, 112
 *-je/o- (pres. stem suff.), 89-90
 **jeug-*, 65, 88, 95, 109, 191; **ju-ne-g-ti*, 191; **ju-n-g-*, 366; **iungenti*, 65; **iung-(e)nt-*, 97; **iugom*, 55, 189, 228, 248
 io-* (H₂io-*), 130, 148, 194, 277, 279, 384; **ios*, 403; **ios-k^wis*, 280
 **iō-*, **iō-*, 120-1, 214, 347
 **ios-*, **is-* (compar. suff.), 121, 122
 **iōs-*, 394; **iōs-nā*, 394
 **kadh-*, 72
 **kaiko-*, 61
 **kai-lo-*, 322
 **kamp-*, 72
 **kan-*, 72, 294
 **kap-mi*, 393, 394, 396, 397
 **kapto-*, 302
 **kata*, **kmta*, 134
 **keh₂-*, 263
 **kel-*, 297
 **kelh₂-*, 72
 **kem*, **kṛṇ*, 135
 **keres*, 263
 **kes-*, 52, 168
 **kjeu-*, 181
 *-ko-, 121
 **kom*, 134
 **krei-*, 241
 **krētus*, 110
 **krenb₂-*, 238; **krenb₂-s*, 112
 **kṛmos*, 113
 **krót-u-s*, 110
 **kuzdho-*, 303
 **kas-*, 72
 **ke*, 129, 135, 168, 279
 **keh₁-*, 397; **keh₁-mi*, 393, 396
 **kei-*, 61, 173; **kei-o*, 86; **kéi-oi*, **kéi-roi*, 192; **keu-ro*, 207
 **kerd-*, 52; **kṛd-*, 162, 228, 248, 365, 380
 **kerh₁-os*, 112
 **kṛ-*, 129, 130, 135, 167, 344, 364
 **kleu-*, 189, 191, 354; **kḷ-né-u-ti*, 191; **ke-kloy-*, 281; **kleu-s*, 365; **kleu-o*, 310; **kleuos*, 360, 368
 **kleuos ṛdḡg^whitom*, 239
 **kṛh₂-to-*, 229
 **kṛntom*, 131, 248, 353, 354, 366
 **kred dbeh₁-*, 98
 **krei-*, **krei-dhro-*, 118
 **kṛh₂-es*, 240
 **kṛh₂-tó-*, 189
 **kuento-*, 386
 **kūo(n)*, **kūō(n)*, 64, 66, 109; **kūn-os*, 380; **kunbhi*, 65
 **k^wa-k^wo-*, 290
 **k^we*, 134, 146-7, 176, 209, 239, 252, 278, 279, 314, 402, 407, 408
 **k^wel(H)-*, 117; **k^wel-eti*, 182; **k^wl-e-*, 396; **k^we-k^wl-o-*, 117, 354; *(e)pi-k^w(e)l-o-, 361
 **k^wetyor-*, **k^wtyor-*, 62; **k^wétyores*, 131; **k^wetyr-sr-*, 120; **k^wetyr-t-*, 132
 **k^wo-*, 52, 130, 134, 194, 387, 397; **k^wos*, 365; **k^wod*, 314; **k^wom*, 393; **k^wi-*, 52, 129, 130, 148, 173, 370; **k^wis*, 161, 168, 170, 230, 354; **k^wid*, 133, 331
 **k^woi-neh₂*, 370
 **k^wrinh₂-*, 359; **k^wri-neh₂-ti*, 197; **k^wrih₂-to-*, 281
 **laiyo-*, 230, 371
 **laku-*, **loku-*, 75
 **lēg-*, 93, 393; **lēg-tó-*, 63
 **lēgh-*, 354, 375; **lēgh-eje-*, 375
 **leh₂-*, 89
 **lēg-*, 393
 **lēgh-*, 341
 **leik^w-*, 230, 366; **lik^w-*, 380; **elik^wet*, 340, 346; **le-loik^w-*, 230
 **leip-*, 189
 **leuk-*, 55, 342, 353
 **liḡ-*, 394
 *-lo- (n. suff.), 117
 **loksos*, 40
 **luHs-*, 355, 362
 **mad-*, 72, 183
 **manu-*, 322, 327
 **māter*, **māter-*, 60, 61, 112, 183, 230, 354, 380, 394, 402
 **me-*, 127; **me*, 129, 358; **med*, 260; **me-ḡh(i)-*, 127; **mene*, 127; **moi*, 129; **eme-*, 127; **eme*, 156
 **mē*, 133, 360
 **med-*, 266, 322
 **medh₂*, 204

- *medh(i)io-, 121, 253, 264, 322
 *medhu, *medbu-, 53, 169, 171, 294, 366, 380
 *megh-, 133, 189
 *mei-, 290; *moino-, 61
 *melh-, *molh-, 164, 263
 *melit-, *melit-, 112, 169, 171, 285
 *men-, 55, 77, 89, 93, 209; *me-món-, me-mn-, 93, 250; *mon-éje-, 249; *mñ-, 89; *men-os-, 209; *mén-ti-s, 108, 113; *mñ-téi-s, 77
 *mezg-, 303
 *mi-, 85
 *-mi- (n. suff.), 210
 *mizdho-, 60, 182
 *-mmo-, 122
 *mñ (n. suff.), 111
 *-mo- (n. and adj. suff.), 117
 *-mo- (ordinal suff.), 132
 *-mon- (n. suff.), 289
 *mōros, 276
 *mñ-, 367; *mñ-to-, 56, 349; *mñ-tro-, 304
 *mūs-, 61, 371, 394
 *nas-, 72; *nās-, 75; *naso-, 370
 *ndher, *ndhos, 134
 *ñ-dhg^hhitom, 228
 *ne, 55, 123, 133, 149, 182; *ñ-, 56, 123, 133, 149, 182, 276, 304; *ne-k^we, 314
 *nebh-es-, 167
 *nehj-, 72
 *nem-, 314
 *nepōt-, 214, 394; *nept-, 24; *nept-ih-, 394
 *nes-, 239, 360; *nos-tos, 230
 *neuy-, 131, 354; *neuy-o-, *neuy^hmo-, 132
 *neyo-, 90, 383; *neu-jo-, 383; *ney-eh^h-je-, 90, 249
 *ni, 134
 *ni-sd-o-, *ni-zd-o-, 55, 63, 73, 188, 248, 303
 *-no- (n. suff.), 117
 *-nó- (v. adj. suff.), 98
 *nok^wt-, 294, 393, 394; *nok^wts, 108, 366; *nek^wt-s, 105, 108
 *nos, 129, 394; *nōs, 395; *ns-, 127; *ns-me-, 127, 128
 *nter, 248
 *nu, 135
 *nu, *nūn, 61
 *-ó- (poss. suff.), 116
 *obho-, 156
 *oi-, 131; *oi-ko-, 131; *oi-no-, 131; *oi-yo-, 131, 357
 *oiso-, 356
 *oktō(u), 131, 183, 188, 302, 304, 370; *oktōy-o-, 132
 *ōkū-, 230, 239
 *ōlenā, 394
 *-on- (individualizing suff.), 305
 *-os, *-es- (s-stem suff.), 112
 *pant-ih^hzm, 355
 *ped-, 51, 63, 108; *pod-mñ, 228, 341
 *peh^h-, 71, 262; *peh^h-s-, 58
 *peh^h-mñ, 108, 111, 162, 341, 348
 *peusk-, 290
 *peku-, *peku-, 207, 262, 386
 *pelekus, 72-3
 *penk^we, 131, 230, 247, 341, 393, 394; *penk^w-to-, 132, 228
 *pent-ob^h-s-, *pnt-h^h-es, 204, 208
 *per-, 325, 404
 *perd-e-, 394
 *perh^h-, *pñh^h-, 132; *pñh^h-yo-, 132, 354; *pñh^h-mo-, 132
 *per(i), 134
 *pet-, 341
 *ph^htēr, *ph^hter-, 64, 112, 188, 189, 204, 230, 275, 276, 304, 342, 344, 354, 355; *ph^h-tēr-s, 109; *ph^h-tros, 344
 *pih^h-uerih^h-, 58
 *pih^h-, 394
 *pi-n-g-, 358
 *pi-ph^h-e-, 249
 *pi-tu-, 290
 *pleh^h-to-, 394
 *pleth^h-, 72, 188; *plth^h-u-, 228
 *pleus-, 395
 *plh^hno-, 98, 189, 304, 367, 375
 *plh^hme^h-, 248
 *plh^hno-, 276
 *potuh^h-, 183
 *prek^h-, 214, 261, 353; *perk-, 294; *prēk-s-, 92; *pñk-ske-, 90, 261, 345
 *pro-, *prō-, *pro-, 55, 132, 134, 282
 *pñt-, 71
 *pu-tlo-, 214, 248
 *-ro (3 pl. middle), 87
 *-ró-, 121
 *rōs-, 395
 *sab-, 406
 *sal-, 60, 72, 75, 342
 *sank-, 249
 *sark-, 21
 *sau-so-, 61, 230
 *sed-, 55, 73, 117; *si-sd-, *sizd-, 228; *sod-éje/o-, 90; *sed-lo-, *sed-leh^h-, 117; see also *ni-sd-o-
 *segh-es-, 188
 *seh^h-uōl, 72, 111
 *sek^w-, 'follow', 263
 *sek^w-, 'see', 314, 393
 *selbh-, 407
 *selk-, 393
 *sem- 'draw water', 290
 *sem- 'one', 131, 132; *smu(i)h^h-, 131, 348; *snñ- (pref.), 132
 *seng^w-h-, 302
 *seno-, 342
 *sepel-je-, 249
 *septmñ, 131, 182, 248, 341; *septm^hmo-, 132, 386; *septm^h-ti-, 393
 *sinos, 394
 *-ske/o- (pres. stem suff.), 90
 *ski-ā, 393
 *smokru-, 348
 *sneh^h-, *(s)neh^h-tlo-, 118
 *sneh^h-, 290
 *sneig^w-h-o-, 365; *snig^w-h-, 228; *snu-n-g^w-h-eti, 253
 *snusos, 60, 113, 116
 *so *seh^h *tod, 127, 130, 306; *so-, 129, 395, 403
 *sol(H)uo-, 398
 *solp-, 393
 *sorb-eje-, 393
 *soro-, 120, 167; *-s(o)r- (fem. suff.), 120; *-sr-, 131
 *spek^h-, 70-1
 *spend-, 162

**speu*-, 302
 **sybh*-, 341
 **sreu*-, 117, 290; **sreu-my*, 61, 230
 **steg*-, 70-1, 267
 **steh*-, 90, 92, 118, 322, 361, 371; **sth*₁-*tó*, 87; **(e)-steh*₂-*t*, 92; **e-sth*₂-*mt*, 233; **sth*₂-*tó*-, 57, 188, 229, 248; **sth*₂-*dhlom*, 118
 **steigh*-, 53, 371; **(s)teigh*-*ó*, 285; **(s)teigh-es(t)*, 285; **stoigh*-*o*-, 395
 **ster*-, 341
 **steu*-, 72, 88
 **(s)tonh*₂-*je*-, 249
 **styh*₁-*tó*-, 248
 **sú*-, 356
 **suHnu*-, 383
 **sub*₂- 'rain', 393; **sub*₂-*énti*, 358
 **sumén*, 64
 **sūs*, 393
 sue*- (se*-), 130, 375, 384; **suo*-, 120, 143, 375
 **sueh*₂-*du*-, 120, 230; **sueh*₂-*ro*-, 353
 **sueks*, 131, 210, 386; **suek(s)*-*to*-, 132
 **suekuro*-, 117; **suekuro*-, 117
 suep*-*no*- (suep*-*no*-, **sup*-*no*-), 118; **sup*-*no*-, 118, 368, 393, 394
 **suesör*, 60, 120; **suesor*-, 207, 341, 342; **suesr*-, 326
 **tar*, 135
 -tāt*- (-teh*₂-*t*), 112
 **tauro*-, 371
 **te*, 129; **toi*, 129
 **telh*₂-, 361; **tłh*₂-*tós*, 229
 **tem*-, 116
 **ter*-, **-tor* (agent n. suffs.), 111
 **-ter*- (kinship n. suff.), 112
 **terh*₂-, 72; **tłh*₂-*é*-, 189
 **terk*^w-, 162
 **-tero*- (compar. suff.), 122
 **ters*-, 52, 182, 254; **trsto*-, 276; **tors-eje*-, 393; **tersä*, 254
 **teuta*, **toutä*, 19, 264, 285, 406
 **-ti*- (n. suff.), 113
 **tkei*-, 182
 **tkei*-, 71, 182

**-tlo*-, 118
 **-to*- (ordinal suff.), 132
 **-tó*- (poss. adj. suff.), 121
 **-tō*- (v. adj. suff.), 98, 164
 **to*-, 129, 344, 361, 395; **tod*, 227
 **-tos* (adv. suff.), 107
 **treb*-, 269
 **trejes*, 123, 131, 341; **t(r)h-sr*- (fem.), 120; **tri-dkmt*-, 279; **tri*- (pref.), 123, 131; **tris*, 133; **tri-t*-, **tł-t*-, 132
 **-tro*-, 118
 **tu*-, **tuH* 128; **tu*-, 156
 **-tu*- (n. suff.), 113
 **tud*-*é-ti*, 89
 -tūt*- (-tuh*₂-*t*), 112
 **uper*-, **uper*-, 134, 314, 341
 **upo*-, 134, 297
 **up*-*s*-, 290
 **uādh*-, 348
 **uai*, 135
 **uē*-, 134, 146-7
 **uedh*-, 366; **uēdh-s-me*, 92
 **uegh*-, 56, 386; **(é)-uēgh-s*-, 92, 257; **ueghst*, 183; **uegh-no*-, 118
 **ueh*₁-, 127
 **ueh*₁-*ro*-, 289
 **ue(t)*-, 127; **uei* (nom. pl.), 127
 **ueid*-, 92, 109, 341; **(é)-uid*-*e*-, 92; **uoid*-, **uid*-, 314; **uoid-h*₂-, 94; **uaduōs*, 98; **uid-to*-, 63, 181, 188, 204, 261, 275; **uidtör*, 230; **uid-tu*-, 275
 **ueik*-, 195; **uik*-, 213
 **uek*^w-, 93, 195; **ue-uk*^w-, 209; **ue-uk*^w-*e*-, 93; **uōk*^w-, 195
 **uel*-, 308; **uelih*₁-*t*, 303
 **uemh*₁-, 72
 **uen*-, 189
 **-uent*- (adj. suff.), 112, 173
 **-uer/n*-, 111
 **uerğ*-, 322; **uorğo*-, 341
 **ues*- 'clothe' 90; **uos-éje/o*-, 90, 394; **ues-nu*-, 345; **ues-my*, 241
 **ues*- 'good' 169, 312; **uesu*-, 209
 **ues-no*-, 118
 **uidheueh*₂-, 261
 **uidhu*-, 290
 **uH-ro*-, 327
 **uikmt*-, 131, 342, 356
 **uis*-, **us*-, **us*-, 61, 355
 **uik*^w-*os*, 56, 74, 116, 182, 366, 367, 393
 **uoh*₂-, 169
 **uobhs-eh*₂-, 74
 **uod*₂-, 111, 162, 165; **ud-or-h*₂-, 106
 **uog*^w-*h-ro*-, 264
 **uor-mo*-, 302
 **uortokos*, 72
 uos*, 129; **us-me*- (us-me*-), 128
 **urēn*-, 394
 **uys*-, 380
 **urt-tó*-, 188
 **urugh-ā*, 404

II. Anatolian

Common Anatolian

**emu*, 156
 **-enti* (erg.), 156, 157
 **obo*-, 156
 **ug*, 156

Hittite

adanna, 164
āki, *akkanzi*, 163
akuwanzi, 172
alpaš, *alpeš/alpa* (OH coll.), 118, 156
ammiyant-, 133
ammuk, 128, 135, 156
-an (obj. pron.), 174
anda, 162, 172, 173, 174
antuhšas, 164-5
anzāš, 162
apā-, 156
appenzan, 129
ar-, 164; *ārḫi*, *arḫa(ha)ri*, 163
arnu-, 164; *arnumi*, 88; *arnummeni*, 161
aršiya-, 171
arunaš, 113-14, 115; *aruna* (OH dir.), 105, 165
-aš (subj. pron.), 129
aši, 176
āššū, 106
atanzi, 172
-ātar, 165
dāḫḫe (OH), 164

dai-, 164
daškmun, 91
daššu-, 162
-e (subj. pron.), 129
ēni-, 130
ēpmi, 163
ēšhar, 111, 112, 165
ešharwant-, 112
ēšmi, 87; *ēšzi*, 58, 60, 161, 163, 164; *ašānt-*, 164
-ēššar, 165
ēzzi, 63; *ēzten* (OH), 63
gēnu, 52
ḡalanta-, 337
ḡalkiś, 164–5
ḡan-, 342
ḡant-, 58, 71, 173, 174
ḡappeššar, 172
ḡarapp-, 58
ḡāraś, 162; *ḡaran-*, 342; *ḡāranan*, 167
ḡark-, 164
ḡarmink-, 164
ḡašduir, 342
ḡaššikke-, 172
ḡaššuš, 164, 333
Ḥatti, 158
ḡuḡḡaś, 342
ḡuiś-, 342
ḡuiśu-, *ḡuiśwant-*, 171
ḡūmant-, 104, 106, 164
ḡupparaś, 104
ḡūwant-, 196
ianzi, *ienzi*, 172
idālawahḡ-, 164
idālu-, 164–5, 347
irḡāizzi, *irḡāitta*, 82
išḡiul, 158
išḡiyanzi, 170
-(i)ške, 164
išpant-, 158
išpantuzzi, 162
ištarniya-, 120
it, 167
nukan, 55
iyant-, 97
iyawanzi, 164
-kan, 135, 239
karaś, 263
kard-, 52, 162
kaś, 161; *kē*, 167
katta, 134
kinun, 135
kiśḡahari, 163
kiśš-, 52
kiśšaraś, *kiśšar-*, 112, 168

-ku (OH), 134
kuenzi, 53, 161; *kunanzi*, 163; *kunānt-*, 97, 162, 164
kugullaś, 117
kuinnaś, 52
kuiś, 130, 161, 170, 176; *kuiit*, 52
kuiśki, 134
kuiś kuiś, 130
kuiit (conj.), 134
laḡḡiyawanzi, 97
lāman, 111; *lamn-*, 90; *lamniyazzi*, 90
lāman dāt, 34
lē, 133
lukke-, 55, 161
luttāi, 361
-ma, 6, 165
marša-, 91; *marše-*, 91
mēk, 133
-mu, 128
nāi, *neyari*, 82; *naiś*, 92
natta, 55
negnaś, 171
nekuz, 105, 108
nēpiś, 167, 169
nešummuk, 158
newa-, 383; *newaḡḡ-*, 90, 249
nu, 135, 157, 165, 176
paḡḡur, 111, 162, 165; *paḡḡuenanza*, 165
paḡḡ-, 58, 71, 262
pāi, 174; *pau*, 157
pāi- 'go', 164, 167; *paiddu*, 95, 157; *pāndu*, 95; *pānt-*, 164
palḡi-, 162
paprātar, 165
parā, 55, 134
pēr, 176; *parnā* (OH), 105
śakki, *šekkanzi*, 163
śallayaś śiunaś śalliś, 122
śarnik-, 21
śāwitar, 172
šeppt-, 112
šēr ar-, 22
šipanti, 162
šiptamiya-, 131
šiunan (OH), 165
-ššaraś, 167
tāwēzzi, 261
taranzi, 171
tarku-, 162
tāru, 52

tekan, 109, 165, 167, 354; *taknaś*, 109; *takni*, 165
teni, *-tani* (OH), 162
terryaś, 131
tuk, 135
tumēnu, 164
ūk, 127 (paradigm), 156
uppeššar, 165
uttar, 104, 164–5
uwai (OH), 135
walḡiśkizzi, 90
walkuwa-, 56
waššezzi, 90
wātar, 111, 162, 165; *widār*, 106, 165
-weni, 161, 162
weriya-, 172
wewakki, 94
zaḡḡiyawenu, *zaḡḡiyawaštati*, 82
zik (paradigm), 128

Luvian (Cuneiform unmarked)

aduna, 170
amu (Hier. Luv.), 156
anda, 173
arkamanallau, 170
-ašša-, 169, 175
attuwat, 347
dām-, 171
dūpaimmi-, 170
ḡant-, 174
ḡapiu, 170
ḡirut-, *ḡiruzaniya-*, 170
ḡiḡḡiyanti, 170
ḡuitwat, 171
inḡagan, 60
iššara-, 168
it, 170
kišammu-, 98, 168, 170
kuiś, 130, 168, 170
kuiś-tar, 135
kuršawar, 156
maddu, 53, 169, 171
malḡaśša, 169; *malḡaśšanzašša-*, 169
mallit, 169, 171; *mallitati*, 169
mammalḡummi, 170
nakkuśiāunta, 170
**nani-*, 171
pa+ra/i- (Hier. Luv.), 169
pāta-, 51

-ša (neut. particle), 169
 -šri- (fem. suff.), 171
tappāš-, 169
-tar, 135
Tarḥunt-, 173
Tarḥuntazalmaš, 35
tatarrya-, 171
tātūš, 176
tipas- (Hier. Luv.), 169
tiwaz, *tiwat-*, 172;
tiwataniya-, 170
tiyamm(i)-, 169; *tiyammašša-*,
 169
tuwi- (Hier. Luv.), 131
uwaland-, 171
wali-la-u-ta (Hier. Luv.),
 170
wāna-, 168
wārša, 169
waššar-, 169
waza- (Hier. Luv.), 56
Wiluša, 168; *Wilušan*, 156
zārza (Hier. Luv.), 52
zi-, 173
ziddāḫiša, 118
ziti-, 118
ziyari, 86, 168

Palaic

-āga- (v. suff.), 171
aḫuwanti, 172
āntienta, 172
apa-, 156
atānti, 172
ḫapittalan, 172
ḫarāš, 171
ḫašan, 172
iška, 90
mušānti, 172
panāganzi, 172
šāwidār, 172
šuwāru, 172
tryaz, 172
uērti, 172

Lycian

abntāi, 173
amu, 156
-ē (obj. pron.), 174
ebe-, 156
kbatra-, 173
kbi-, 173
kumaza-, 174

mahāna, 176
Natr-, 174
rite, *rite-*, 173, 174
pddē, 173-4
pijata-, 118
pije-, 174
st-, 173
tadi, 157, 169
tede-, 176
θē, 173, 174
ti-, 173
-ti (refl. part.), 174
tudem, 174
Trqqas, *Trqqnt-*, 173
-wāt(i)-, 173
wawa-, 60, 173
ḫnt-, 173
ḫntawati, 174
ḫssaθrapazate, 174

Lydian (and other)

ak, 176
alarmš, 175
amu, 156
bidv, 175
biraλ, 176
est, 176
ētamv, 175
fak, 176
ist, 176
-k 'and', 176
klīdal, 176
koful, 176, 347
mašara (Sidetic), 177
nāqis, 176
-si- (adj. suff.), 175
tašev, 174
red (Carian), 177
-xi (Carian), 148, 177

III. Indo-Iranian

Common Indo-Iranian

**ārya-*, 19, 187
 **m̥as dhā-*, 210
 **nām* (gen. pl.), 183
 **urvarā-*, 209
 **yā-* (passive), 183

Indic

**atka-*, 184

MITANNI INDIC

a-i-ka-, 184, 190
a-i-ka-wa-ar-ta-an-na, 184
in-dar, 184
mi-it-ra-, 184
na-a-wa-ar-ta-an-na, 184
na-ša-at-ti-ya-, 184
pa-an-za-wa-ar-ta-an-na, 184
u-ru-wa-na-, 184

SANSKRIT

ābravīt, 189
adán, 193
addhā, 181, 204
ádharma-, 209
adhás, 134
(á)dhāt, 157
agnis, 195
ahám, 127
āhann āhum, 210
ānuṣṭan, 93
ākṣi-, 210
āmīti, 241
amfta-, 207
a(n)-, 56, 133, 182
ānaikṣit, 190
antár, 134, 214
āpa, 134
Apām Nápāt, 24
apás 'waters', 209
āpas-, 110
apás-, 110
apāśca urvárāśca, 209
āpi, 134
āpiparas, 93
āpnōti, 347
ārcati, 360
ārya-, 187
āsmu, 87; *āsi*, 63; *āsti*, 60,
 161, 195; *sānti*, 161; *ās*,
āsit 195, 214; *āsat*, *āsati*,
 207, 250; *syām*, 96; *syās*,
 261; *sat-*, 97
asmin, 129, 268
aṣṭā, 183, 188
aṣṭamā-, 132
āstoṣi, 192
āsú-, 239
āsura, 333
atiṣṭīgham, 53
atmā, 150
ātsi, 63; *attá*, 63
āndat, 92

ayám, idám, 130
āyu, 108
babhrū-, 112, 184
bahuvrīhī-, 122–3
bālam, 52
barhīs, 42
bhāgas, 112
bhāgavant-, 112
bhārāmi, 53, 89; bhāre,
 85–6; *ābharam, 85, 91;*
 ābhari, 85–6; bhāreyam,
 96; *bhāra, bhārata, 95;*
 bhāranti-, 97; bhāramāṇa-,
 97
bhinná-, 98
bhūta(na), 7
bódhati, 188; buddha-, 182,
 188
bráhman-, 110, 195
brahmān-, 110, 195
br̥hānt-, 182
ca, 134, 239, 261
cakāra, cakāra, cakr-, 181,
 183
cakrām, 36, 38, 117
cāratī, 182
cātasras, 120, 131
cātuṣpad-, 131
cyavante, 181
dā-, 257; dādāti, 88, 207;
 dādati, 88; dadáu, 256;
 dāsyāti, 91; dātum, 97;
 dātā, 112; dātā, 111
dāhate, 354
dāmiya-, 121
dāmpatis, 108, 207
dāru, 52; drós, 108
dāsa, 56, 181; daśamā-, 132;
 daśāt-, 112
dātā vásūnām, 112
dātā vásūni, 112
devās, 22, 115–16, 117, 190,
 193, 204, 205, 210; *devān*
 (devāms), 116, 194
Devāśravās, 35
devī, 120
dhā-, 188, 341; dādhāmi,
 195; *dādhati, 207*
dhāma, 21
dharāyati, 214
dhīṣṇya-, 57
dhvan-, 209
dirghā-, 182
ditī-, 57
duhitār-, 204
durñāman-, 188

dvēsmi, 191–2
dvipād-, 131
dvīs, 133
dyaús, dryāus, 61, 66, 109;
 dyām, 64
dyāus pitar, 22
éka-, 131, 184, 190
eśām, 129
étu, 157
gam-, 262; gatā-, 98
gāus, gāv-, 60, 181, 190;
 gām, 64
gāuya-, 121
(g)ha, ghā, 135
gharmā-, 117
ghōśas, 181
gnā-, 207
gopās (paradigm), 104
grāsate, 52
grāvān-, 111
gr̥bhñā-, 207; gr̥bhñāmi,
 gr̥bhñāti, 191–2
gurū-, 111, 113
ha, 374
hānti, 53, 188; hatā-, 98;
 jīghāṁsati, 91; hatvā, 192
hāri-, 207
hastī, hástī, 104
hitā-, 188
hūyāte, 53
ihī, 167, 188
jagāma, 94
jan-, 195
jānas-, 112
jāni-, 52, 119
jānu, 52
jigāti, 88
juṣṣa, 93
jyā, 116
kalpāyati, 195
kām, 135
kāmas, 263
kārma, 111
kās, 52, 130, 134, 194
kās ca, 134
kavī-, 31
kim, 194
krātus, 110
krāviṣ-, 112
kr̥nāti, 197, 285; kr̥tā-, 197
kriyānte, 183
kṛnavā, 96
kṛṣṇa-, 120
kṛṣṇā-, 120
kṛtā-, 207
kṣām, 109, 209; jṁās, 109

kṣāratī, 182, 197
kṣatram, 213
kṣāyati, 182, 214
kṣétram, 182
kúba, 107
liptā-, 189
mā, 133, 361
madati, 183
mādhū, 53
māhi, 133, 189
-mahī, 204
māhyam, 128, 261
mānas, mānas-, 55, 195,
 209
mántra-, 207
mānu-, 322
māryate, 89
mártya-, 207, 214
mātār-, 60
matīs, matī-, 108, 110, 113;
 matī, 105
me, 128
medhā-, 210
meghā-, 341
mēh-, 71
miḍhām, 182
mriyāte, 345
mṛtā-, 56
mugdhā-, 181
muṣṇāti, 358
nā, 55, 133, 182
nābhas-, 167
nadínām, 183
nāma, 111; nāman, 105
nāma dhā-, 34
nāmas, 207
nār-, 31, 71
naṣṭā-, 182
nāva, 184
nāveda-, 7–8
nāvya-, 383
nī, 134
niḍās, 55, 188
nū, 135
nunām, 61
ojas-, 210
ótum, 192
pā-, 262; pātave, 97
pād, pad-, 51, 104–7; pattās,
 patsutās, 107
palitā-, 184
pāñca, 184
pāri, 134
pāsu, 262
paśumāntī, 106
pāsyati, 70

pibati, 294; *pība*, 269;
pībadhya, 97
piṅgalā-, 184
puta, *putār*-, 188, 189, 193
pñtaye, 97
pñtu-, 290
pñvarī, 58
prā, 55, 134
pra math-, 27
prākṛtam, 189, 196
prāthimān-, 188
prāti, 241
prcchāmi, 208, 209; *prcchāti*,
 90
priyā, 119, 193
prñāti, 188
punāti, 75–6, 88; *punimās*,
 88
purās, 189
pūrñā-, 98, 189
pūrviyā-, 132
putrās, 214
rājān-, 322
rāt, 60
ratheshthā-, 123
ravis, 348
rāyasi, 89
rēriha-, 117
rghāvān, 30–1
riptā-, 189
rjā-, 121
rjū-, 209
rk, 360
rócate, 55, 93; *ruróca*, 93,
 94
rtām, 31
rtāsya pathā, 205
rtāvān, 30–1
rtvij-, 109
rudhitrā-, 71, 121
sā sā tād, 130, 194, 209;
tāsmā, 385
sācā, 214
sāda, 55
sābas-, 188
sahāsra-, 210
sakṛt, 132
salilā-, 60
saṁdhī-, 190
saṁgṛbhya, 192
śāmsa-, 31
saṁskṛtam, 189, 196
saptā, 182
śastā-, 31
śāt, 210
śatām, 353; *śatatamā*-, 132

śāye, 86, 168, 192; *śére*, 87,
 192, 207
Śmhapuras, 122
sindhuka-, 121
śiras, 240
śirtā-, 189
ślōkas, 189
śmāśru-, 348
snuśā, 60
sōmas, 190, 204
spās-, 71
spāśāyate, 70
śrād dadhāti, 98; *śrāddhā*-,
 52
śramaṇas, 352
śrāvas ākṣiti, 29
śṛṇōmi, 191–2; *śṛṇōti*, 191;
śṛṇvé (3rd sg.), *śṛṇviré*,
 192; *śrāvi*, 192; *śrudhī*, 95
sthā-, 257; *āsthām*, 192;
āsthāt, 92; *tasthāu*, 257;
sthānt-, 97; *sthitā*-, 57,
 188; *sthāpāyati*, 195
su-, 71, 123, 214
śūcis, 193
sūnāra-, 71, 189
śūnas, 380; *ś(u)vā*, 66, 109
svā-, 130
svādsi-, 120
svāsā, 60
svāśura-, 117
śvāśura-, 117
svayām, 130
śyāmā-, 383
-tama- (compar. suff.), 122
-tama- (ordinal suff.), 132
tanōti, 88, 191; *tanvānti*,
 88
-tara-, 122
tarśāyati, 52
tirāte, 189
tisrās, 120, 131
tripād-, 131
trīs, 133
tṛṣṇā, 182
tṛṣṭā-, 276
tṛṭīya-, 132
tudāt, 72, 89
tvām, 128, 194; *tvā*, 194,
 209
tyājas-, 206
u, 134
ukṣā, 109
ūpa-, 134
upāri, 134
urvi, 105

Uśās, 23; *uśas*, 62; *uśāsam*,
 109, 110; *uśās*, 62, 109
uśnā-, 188
ūtibhis, 107
vā, 134
vac, 195; *āvocam*, 93, 191;
voca, 209; *vōcam*, 191
vācā, 105
vāhati, 56; *āvākṣam*, 192;
āvāt, 183; (ā)vakṣur, 92;
vakṣayam, 182
vaiśvānarā-, 65
vāśīyah, 195
vām, 194
vanōti, 189
vār, 169
vartanam, 184
vas, 194
vāsāyati, 90
vāsu-, 209, 214, 312
vasuśravas-, 406
vātas, 189, 195
vāti, 196
vāyati, 375
vāyau īndraś ca, 140
vāyūs, 195
vēda, 94, 341; *vidūṣ*-, 98
vēdas, 7, 185
vidmāne, 97
vīndat, 191; *āvīndat*, 191
vīrás, 113–14
viṣ-, 61
viśva-, 383
viittā-, 181, 188, 204
vṛkas, 56, 182, 207; *vṛkīs*, 120
vṛtrahā, *Vṛtraghṇ*-, 109, 347
vṛttā-, 188
yā-, 130, 194
yājate, 228
yajñā-, 117
yātate, 360
yugām, 55, 189; *yugā* (du.),
 115
yunājmi, 191–2; *yunākti*,
 75–6, 88, 191; *yunājanti*,
 88; *yunājati*, 120
yūvā, 111
yuvām, 194
yuvāti, 383
yūyām, 194

PĀLI

kiṇati, 197
paggharati, 197
uggharati, 197

HINDI

pañj, 131

Iranian

*anwa-, 184
 *arya-, 187
 *baga-, 368
 *dānu-, 368
 *druj-, *drug-, 210
 -ka- (n. suff.), 214

AVESTAN

a-, 182
 aḍairi, 134
 aḍara-, 209
 adā (OAv.), 209
 aēnuu-, 131
 āfš-, āp-, 182; apō-, 209
 ah-, 207; ahmi-, 204, 208;
 asti, 205; ās-, 214; aṇhaṭ-,
 207
 ahura-, 205
 aburō mazdā-, 205
 aiuarā, aiṭan, 208, 322; aiiārō
 (OAv.), 106
 amaša-, 207
 anusant-, 149
 aoyā (OAv.), 204; aogəda
 (OAv.), 181, 204; aoxta-,
 204
 aojah-, 210
 Apam Napā-, 24
 aš-, 210
 aša-, 210
 ašahiā paθō-, 205
 aši-, 210
 aspaēšu-, 116
 aspōgar-, 206, 207
 at-, ad-, 208
 auua-, 214
 auuapastōiš (OAv.), 209
 azəm-, 127, 213, 214
 barainti-, 203; barəm-, 91;
 barōiš-, 96; barəmina-, 97
 bərəzant-, 182
 biš-, 133
 brātār-, 182
 ca-, 134
 carauti-, 182
 caθru-gaoša-, 131–2
 caθunarō-, 204; catagrō-, 120
 caxr- (pf.), 181

caš-, 129
 dā 'gift' (OAv.), 208
 daḍāti-, 207; dā (OAv.), 208
 daēna-, 119; daēnam-, 207
 daēnuu-, 115–16, 117, 204,
 205; daēnuō-, daenuas-,
 207, 208; daēnuuəng
 (OAv.), 207; daēnuuēibiš-,
 206
 dāiš-, 61
 dām (OAv.), 105
 dāmiš (OAv.), 210
 darəga- (OAv.), 182
 dasa-, 56, 181; dasəma-, 132
 dāuru-, 52
 dačiū- (OAv.), 214
 dəng paiti- (OAv.), 108, 207
 dərətā-, 209
 dərəzi-raθa-, 123
 dərəzra-, 123
 didrayzō.duiie (OAv.), 206
 duuapuman-, duumman-, 209
 duxtar-, 204
 ərəš-, 209
 ərəzatəm-, 39
 ərəzu-, 209
 fra-, 55, 204
 fraḍda-, 63
 fras-, 214
 fšuiias-, 207
 gaoša-, 181
 gāḇā-, 205
 gāuš-, 61, 181
 γ(ə)nā-, 207
 gərəβnā-, 207
 γzar-, 182, 197
 haoma-, 190, 204
 hapta-, 182, 204
 hazayrā-, 210
 hō hā taṭ-, 130
 hu-, 214
 im-, 130
 iθiiejah-, 206
 jaini-, 52
 jantti-, 53
 janaṭ aṭim-, 210
 kamərəda-, 210
 kərəfš-, 104, 106
 karata-, 207
 karatāe-, 97
 kirijente-, 183
 kō-, 52, 130
 kudā-, 107
 maēθana-, 404
 -maiði-, 204
 manah-, 55

mašua-, 207, 214; mašiiəng-,
 115
 mat-, 214
 mātār-, 183
 maθra-, 207
 mazdā-, 209
 mərta-, 56
 mižda-, 60, 182
 mōi-, 209
 mruue-, 192
 na-, 55
 nā-, 209; nar-, 104; nərəš-,
 208; narəṭ-, 208
 nāidudāgham-, 121
 našta-, 182
 nəmō-, 207
 pairi-, 214
 paṇtā-, 208; paθō-, 204
 pasu vira-, 267
 paθni-, 183
 pərəsā (OAv.), 208, 209
 pitā-, OAv. pitā-, 204
 pitu-, 290
 puθra-, 214
 puxša-, 132
 raocaiieiti-, 55
 rərema-, 117
 satəm-, 53
 šiauu-ai-, 181
 sōire-, 87, 192, 207
 šoiθra-, 182
 spasiueiti-, 70
 spənta-, 386
 surunuuatascā
 asurunuuatascā, 33
 taršna-, 182
 taṭ-, 209
 θā-, 209
 θri-, 210
 θriš-, 133
 θruzafanəm-, 210
 tišrō-, 120
 tū-, 128
 uruuaēsa-, 184
 uruuarāš(cā), 209
 va-, 140
 vaβžaka-, 74
 vaiuθi-, 135
 vaṇhu-, vohu-, 209
 vaocā-, 209
 vaxšaṭ (OAv.), 182
 vazaiti-, 56
 vəhrka-, 56, 182, 207
 Vərəθrayna-, 347
 viθnuā (declension), 98
 viš-, 61

vis-, 213
vista-, 181, 204
xrāra-, 204
xšaθra-, 213
xšunaš, 210
x'ai, 130
x'aphar-, 207
x'ang (OAv.), 111, 207
ya-, 130
yākara, 108
yāgham, 207
yaojā (OAv.), 96
yasna-, 113–14, 117, 205
zafare, 210
zavri-, 207
zam, 209
zaotar-, 53
Zarathuštra-, 205
žnum, 52

OTHER (OLD PERSIAN UNMARKED)

abestāg u zand (Mid. Pers.), 204
adam, 213, 214
āha, 214
apariyāya, 214
Armuniya-, 337
asa, 212
aspa (Median), 212, 214
atar, 214
Auramazdā, 205
avam, 214
azdā, 181, 204
bar-, 214; *bara(n)tiy*, 183, 203
brātar-, 182
cartanaiy, 97
caxr- (pf.), 181
dahyūnām, 214
dauva-, 204
dāraya-, 214
Dārayavaus, 214
fra-, 204
frasta-, 214
fraθ-, 214
hačāma, 214
hašiyam nary duruxtam, 33
hauma-, 204
Haxamanišiya-, 121, 214
haya-štaya-, 213
mard (Mod. Pers.), 349
maritiya-, 182, 214
napā, 214
Ohrmazd (Mod. Pers.), 205

pari, 214
Parθava, 216
puça, 214
šāh (Mod. Pers.), 214
spāka, 368
θātry, 393
ubartam, 214
Varθagan (Parthian), 347
vau-, 214
viθ-, 213
xšaça-, 213
xšaθrapavā, 174
Xšayārša, 35
xšāyaθiya-, 182, 214
-(y)e (Mod. Pers.), 213

IV. Greek (Attic unmarked)

adelpbós, 132
akontos emeio (Hom.), 149
Agélaios, 35
ag-, 342
ainumai, 20
aión, 239
aiópolis, 64
akoúo, 61, 74, 230, 315
Albanoi, 405
Aléksandros, 35
állos, 130
āmar (Doric), 342
āmmes (Aeolic), 237
amphi-, 239
amustí, 172
a(n)-, 56, 133
anáthēma, 58
-anein, 345
anén, 57, 229, 402; *aner-*, 71
ámōge, 93
ánthrōpos, 240; *anthrōpō*, 115; *anthrōpoi*, 230
anti, 134, 325
áphthiton, 228
apó, 134
Ápollōna, 238
ára, *rhá*, 135
argós, 121
Aristódēmos, 123
áriston, 322
árotton, 118, 342
ároura, 209
asámnthos, 222
a-to-ro-qo (Myc.), 240
au, 134
aúlis, 342
aiós (Aeolic), 109
ai(w)éksomai, 164
ái(w)esa, 342
ái(w)esi, 196
básis, 229
(-)ibatós, 98
beltiōn, 52
biós, 116
bōn, 64; *bó(w)es*, 60
boukólos, 64
brotós, 56
dákno, *déksomai*, 91
dé, 135, 151
-de (adv. suff.), 239
dēdae, 93
déd(w)imen, 342
dēka, 56
deksiterós, 122
dēmóboros, 116
dérōn, 239
despótēs, 108, 207
diā thedōn (Hom.), 122
Dinoklēs, 35
Dimokrátēs, 35
dídōmi, 88; *dídomai*, 231; *edídōmēn*, 232; *edómēn*, 232–3
Diogénēs, 35, 400
dípod-, 131
dís, 133
Diweiphilos (Cypriot), 240
dō, 108
dólikhos, 120
dolikhós, 120
dōron, 58, 342
dórpon, 394
dóru, 52, 108
dósis, 57
dotēr, 111
dōtōr, 112
dotós, 229
dowenai (Cypriot), 97, 279
d(w)áros (Doric), 342
ē, 348
edáēn, 233
édramon, 361
egó(m), 127; *egóge*, *emége*, 135; *meu*, 129; *moí*, 128
emí, 87; *eī*, 63; *estí*, 58, 60; *eimen*, 233; *eiēn*, 96, 233
ekēinos, 135
ekselawmoia (Arcadian), 96
élekse, 92
éluthe (Hom.), 92
emoí, 128

- en*, 134
endo-, 173
enkhesimōros, 276
éntheos, 123
énuma (Laconian), 111
Ēós, 23
epí, *epi-*, 134, 341, 408
épleto, 396
érebos, 229, 342, 348
e-re-e (Myc.), 234
erruā (Doric), 250
eruthrós, 71, 121
esthiō, 374
etímēthēn, *tīmēthēsomai*, 233
eu-, 123
Eukratēs, 35
eupátōr, 73, 123
Eurukratēs, 35
é(w)idon, *é(w)ide*, 92, 228
gálakt-, 37
génesis, 229
génos, 112, 227; *géneos*, 105;
génea, 106
gérōn, 342
gígnomai, 88, 90; (-)gnētos,
 229
glak-, 37
glōtta, (Ionic) *glōssa*, 229
gnōmē, 117
gnōtós, 276; (-)gnōtos, 229
gónu, 52, 108, 230
grabos (Mod. Gl.), 266
grástis, 52
guné, 52, 119, 375
hādús (Doric), 120, 230
hágios, 230
hál-, 60
háploos, 132
haūs, 61, 230
hé, 130
hédra, 73
hēgemōn, 289
heís, *mía*, 131, 348
béks, 8; *bekios*, 132
hellá, 117
hemeís (Ionic), 237
hénnumi, 345
heós, 130
hēpar, 108
heptá, 8, 227; *hēbdomos*, 132
hēzomai, 55
hippikós, 121
híppos, 240; *híppō*, 235;
híppōi, 61
hístāmi (Doric), 92; *éstēn*,
 232-3; *éstā* (Doric), 92;
éstan (Doric), 233; *stēomen*
 (Hom.), 233; *stant-*, 97;
statós, 57, 229
hízō, 228
ho hē tó, 130, 194, 227
bó- (rel. pron.), 130
hoi (dat. 3sg.), 403
hōrā, 234
bōstis, 280
hūdōr, 106, 111
hūiei, 393
humén, 64
hupér, 134, 314
húpnos, 118
hupó, 134
hupódra, 227
hupsélós, 290
ídmenai (Hom.), 97
ikbithūs, 61, 230
iós, 230
íphi, 107, 234
-isto-, 306
íthi, 95, 167, 233; *ítō*, 403
-ize/o- (denom. suff.), 90
kakón, 403
kaléo (Hom.), 91
káluks, 240
kámatos, 229
kardiā, 52, 228, 276; *kradiē*
 (Hom.), 228
karēat-, *krāat-* (Hom.), 240;
ka-ra-a-pi (Myc.), 240
káruon, 240
katá, 134
katarwos (Arcadian), 170
keímai, 61; *keitoi*
 (Arc.-Cypr.), 232
kekharisménos, 404
Keltoí, 277
ke(n), 135, 239
kephalé, 354
ke-re-a₂ (Myc.), 227
keskéon, 52
khamai, 105
kheír, *kheir-*, 112, 168
khé(w)ō, 53, 263, 354
khlōrós, 402
khtbōn, 397
kís (Thessalian), 227
kléos áphthiton, 29
-klé(w)ēs, *-klēs*, 112
klé(w)os, 112, 310, 360, 368
klōps, 77, 104
Kórinthos, 222
kratús, 110
kréas, 112
krínein, 241
kruóeis, 238
kūdánewa, 123
kūdrós, 123
kúklos, 36, 38, 117, 354
kuōn, 66, 109, 235; *kunós*,
 380
labúrnthos, 222
lai(w)ós, 230
lákkos, 75
leipō, 230; *lumpánō*, 231;
élipon, 232; *élpe*, 342,
 346; *léloipa*, 94, 230;
lumpánein, 345
lékhos, 354
lektós, 63
leukós, 55
Libukós, 121
litas t' epaoidás, 33
loigós, 393
lúkos, 234
lúō, *lúsō*, 233; *élusa*,
elúsamēn, 232-3
maínetai, 89
manthánō, *mathēsomai*, 82
mátēr (Doric), 60, 230, 239
mē, 133, 342, 361
méga, 133, 189
mégistos, 122
mélitta, 229
mémone, 94
ménos, 55, 195
mērós, *mēroí/mēra*, 118
mésos, 121, 322
mētēr, 230
méthui, 53, 294
me-zo-e (Myc.), 121
minnēskō, 231
misthós, 60, 182
mūs, 61
naūphi, 234
néktar, 29
nēmo, 358; *németai*, 20
nēomai, 239
nēphos, 167
Néstōr, 239
niph-, 112, 228
niphó(w)ent-, 112
nómos, 77
nóstos, 230, 239, 361
nú, 135
nū(n), 61
nūnū, 85
nuós, 60
odelons (Cretan), 115
odúnē, 347

- ógdo(w)os, 132
 oínos, 131
 okruōeis, 238
 ókús, 230, 239
 óllumi, ólōla, 93, 94
 omēikh-, 71
 omnūmi, 241
 óneidos, 345
 ónoma, ónuma, 111, 402
 onomainō, 90
 onomáklutos, 34, 360
 ónoma títhesthai, 34
 óphelos, 229
 ophrās, 405
 ópsomai, 353
 órnith-, 167
 órnūmi, 88
 orthós, 290
 ósse, 105
 ou, ouk(í), 133, 239, 331
 ouranóthi, 107
 palámē, 248
 pánt-, 355
 patrēr, patr-, 73, 104, 109,
 123, 230, 235; páter, 193
 pénte, 230; pémpotos, 132,
 228
 perí, 134, 214
 phégós, 40
 phēmi, 7, 231; phāmi, phāsi
 (Doric), 231; phēsi, phāsi
 (Doric), 239; phantí
 (Doric), 231; éphēn, 232;
 éphān éphās éphā (Doric),
 232; ephānēn, 233;
 (-)phatós, 98
 pheréokos, 123
 phérō, 53, 89, 231; phérousi,
 229; phéronti (Doric), 229,
 231; phéromai, 85–6, 231;
 phérō (1sg. subj.), 96;
 épheron, 85, 91, 228;
 éphere, 227; epherómēn,
 85–6; phérōmen, 233;
 phéroumi, 96; phéroumen,
 233; phére, 95, 233;
 pherétō, 95; phérete, 95,
 233; phéreín, 90; phéront-,
 97; phérousa, 120;
 pherómenos, 97–8
 pheúgeskon (Hom.), 90
 Philippos, 405
 -phomos, 53
 phoréō, 90
 platús, 228, 276
 poiné, 370
 pólis, póli, 104
 poñ, 130
 pouís, pod-, 51, 104–6, 341;
 póda, 228; póde, 235;
 pódes, 342; po-pi (Myc.),
 106
 pró, pro-, 7, 55, 134
 prógnōsis, 229
 Promētheús, 27
 prophētēs, 7
 prós, 241
 protí (Hom.), 241
 pseudēs, 110
 pseudos, 110
 pūr, 341
 -qe (Myc.), 240
 -qe-to-ro-po-pi (Myc.), 227
 rheūma, 61, 230
 Rhódippos, 35
 seu, 129
 seúetai, seútai, 78
 skaiós, 122
 skiā, 393
 Sophoklēs, 112
 sophóteros, 122
 sōtēr, 234
 spéndō, 162
 stégō, 70
 steikhō, 53
 stēthessin (Hom.), 237
 súzug-, 109
 tar (Hom.), 135
 tà zōia trékhe, 118, 235
 te, 134, 240, 261
 Telamónios, 121
 téllomai, e-pl-ómēn, 227
 -tero-, 122
 térsetai, 52
 téssares, 62, 227, 237; písures
 (Hom.), 62, 237; péssures
 (Lesbian), 227
 tétartos, 132
 tetrapod-, 132
 te-tu-ko-wo-a (Myc.), 234
 thánaton, 405
 thēma, 111
 thēmus, 21, 210
 Themistoklēs, 35
 theós, 1, 22, 78, 114–16;
 theā, 119, 239
 thermós, 117, 302, 341
 *thēs-, *thes-o-, 78
 thésphatos, 57, 78
 thugáter, 342
 tís, 129, 130, 227, 230; tí, 52
 tís tar (Hom.), 135
 tībēmi, 195, 227; tithent-,
 97; (é)thēke, 403; thetós,
 229
 (-)tlatos (Doric), 229
 toikhos, 343
 Tókharoi, 351
 tomā (Doric), 118
 tómos, 116
 tomós, 116
 tripod-, 131
 trís, 133
 trítos, 132
 tí (Doric), 128
 tiúkos, 72
 (w)ē, 134
 (w)eidós, 98
 (w)eiþon, 93
 wekhetō (Pamphylian), 56
 wēma, wēmas (Cretan), 241
 (w)épos, 195
 (w)érgon, 322
 ube (Pamphylian), 130
 Wālon, 168
 (w)istōr, 230
 (w)oiða, 94, 228, 232; isanti
 (Doric), 232; (w)istós, 63
 (w)oikos, 195, 228; woikos
 (Thessalian), 228; oíkoι
 (nom. pl.), 62; oíkoι (loc.),
 62, 234
 (w)oiños, 38
 Zeū páter, 22, 261
 Zeús, 61, 66, 109; Zēn
 (Hom.), 64
 zō, 361
 zōnē, 394
 zugón, 55, 228

V. Italic

Latin

- ab, 134
 *ab-oculus (Vulg. Lat.), 279
 absent-, 97
 aeger, 356
 aetate, 6
 affectus, 254
 agō, 89, 342, 353; agimus,
 358; age-, 250; égit, 250;
 agitō, 265; agere, 249, 250
 ait, 348
 albēre, 249
 albēscere, 249
 alius, 130

- alter*, 248
altius, 353
alumnus, 98, 251, 322
amāre, 253, 256
ambi-, 397
amicus, 254
Anglī, 316
angustus, 253
anima, 248, 362
annus, 267
ante, 58, 325
aptus, 254
aqua, 403
arāre, 383
arātrum, 118
argentum, 39, 263; *argentō*, 268
armiger, 123
artifex, 254
atque, 265, 403
attigat (Archaic), 286, 358
audeō, *ausus sum*, 83
audire, 256; *audiui*, 256
augeō, 353
aurōra, 23, 109, 380
autem, 134
barbātus, 121
bellicus, 121
bellum, Archaic *duellum*, 253, 255
bene, 252
bibere, 249; *bibit*, 294; *bibe*, 269
bis, 133
bouēs, 60; *būbus*, 106
caecus, 61
caelitus, 107
calēscere, 394
cambiāre (Late Lat.), 279
canō, 294
capere, 256; *captus*, 251
cārae (dat.), 263
carmen, 111
carniuorus, 116
Castorei (Archaic), 254
Catō, 305
catus, 305
cēlāre, 297
centum, 53, 248, 259, 366
Cerēs, 112
Chattī, 302
cis, 130
colligo, 254
coniug-, 109
cord-, 52, 248
cornū, 255
cōsentiant (Archaic), 256
crēdō, 98; *crēdere*, 296
cribrum, 118, 241
cum (conj.), 134
cum (prep.), 134
cūra, 249
cūrāre, 249
currus, 255
datiō, 57
dē, 347
dea, 119, 239
dēbilis, 52, 347
decem, 56, 304
decimus, 132
dediit, Archaic *dedet*, 250, 269
dēfendit, 53
deiuōs (Archaic), 117
dēlendus, 251
dēpilis, 120
dēplūmis, 123
dēterior, 122
deus, 1, 22
dexter, 265
diēs, 255
dīuit-, 170
dō, 263; *datōd* (Archaic), 95; *dūt* (Archaic), 263; *datus*, 248
domus, 261; *domī* (loc.), 256
dōnum, 118
dūcere, 253, 256; *dūcam*, -ās, -at, 254
duident-, 131
dūrus, 239
ēbrius, 172
ego, 127, 407; *mihi*, 261; *mēd* (Archaic), 263
ēgregius, 121
emom (Archaic), 261
etate (inscr.), 6
ex, 177
facere, 403, 406; *faxō* (Archaic), 250; *factus*, 254
fāma, 117
fānum, 57, 239
far, 263
fātus, 239
femur, 111, 251
ferō, 53; *ferāt* (Archaic), 247
fēstus, 239
fidēs, 20
findere, 249; *fissus*, 248
fingō, 263
foedus, 20, 254
formus, 117, 302
fortissimus, 122
fraud-, 266
fructus, 267
fugere, 249
fumus, 209
fundō, 263
fuīs, 53
Galatae, 277
Gallī, 277
genū, 52
genus, 112, 253; *genera*, 106, 253
Germāni, 300
Gnaiuōd (Archaic), 114, 253
(g)nātus, 248
gnōui, 257
grāmen, 52
grānum, 276, 380
habēre, 256, 258; *habēas*, 251
hic, 135
hoc, 258
homō, 22, 64, 110, 383, 402
hortus, 264, 322
idem, 269
ignis, 195
ilicō, 254
imprūdēs, 149
in, in-, Archaic *en, en-*, 56, 133, 134, 254
inēptus, 254
īnfimus, 247
īnfra, 134
inimicus, 254
inter, 134, 214, 248
inuītus, 149
iouesat, 253
iouxmenta, 254
īre, 249; *it*, 170; *eunt*, 172; *ēitur* (Archaic), 256; *ītō*, 403
is ea id, 130, 269
iugum, 55, 248
iungūt, iungunt, 88; *ungent-*, 97
Iūpter, Iou-, 22, 253, 261, 406; *Diouos* (Archaic), 253
rūs, 21
lact-, 37
lacus, 75
lāpsus, 268
legō, 254; *lēgī*, 93, 396
lex, 21
Liber, 263, 400
libertāt-, 112
ligāre, 394

lingua, Archaic *dingua*, 248
locus, *loci/loca*, 102, 118, 254
lūcus, 383
lupus, 255
lux, *lūc-*, 55, 353
magis, 122; *maior*, *maius*, 122, 253; *maior* (Archaic), 122
magistrātūd (Archaic), 252
malos (Archaic), 254
māne, 261
maneō, 239
manna, 290
manūs (acc. pl.), 264
Mārtei (Archaic), 105
māter, 60
maximus, 122
mēcūm, 139
medius, 121, 253, 264, 304, 322
mē insciente, 149
memini, 94, 250
mēns, 55
meos sokiois (Archaic), 254
meritō, 133
minimus, 122
modus, 322
moenia, 254
molitum, 263
mollire, 249
mollis, 248, 249
monere, 249
morbōs uisōs inuisōsque, 33
moriōr, 345
mort-, 367
Morta, 56
mortuus, 407
mūnus, Archaic *moenus*, 61, 290
mūs, *mūrēs*, 61, 253
nanciōr (Archaic), 256
nāre, 290
nārēs, 75
nē, 133
ne-, 55, 133
nebula, 253
nefas, 133
Neptūnus, 24
neque, 314
Nerō, 267
nescire, 55
neu, 256
nidus, 55, 248
ninguit, 253
nix, *niu-*, 248

nōmen, 111, 248, 267
nōmen indere, 34
nōnus, 132
nōs, 127
nouāre, 249
nouus, 383
nox, 108
nunc, 135
nūper, 61
nurus, 60
ob, 134
ob uōs sacrō (Archaic), 140
ocior, 239
octāuus, 132
oculus, 315, 348, 361
odor, 342
onus, 254
onustus, 254
opus, 269
oquoltōd (Archaic), 254
orbis, 58
ōs, 322
palma, 248
parēns, 263
pariō, 263
pāstor, 58
pāstōrēs pecuaque, 262
pater, 23, 276; *patri*, 254; *patre*, 6
pater familiās, 23, 119, 255
pecū, 19, 207
ped-, 51
per, 134
piāculum, 267
Pictī, 291
pingunt, 358
piscis, 290
plānum, 276
plēnus, 98
-plētus, 98
pōculum, 248
poena, 257
Popliosio Valesiosio (Archaic), 114, 252
populi, 264
porculus, 117
poscō, 261, 345; *poscit*, 90
precor, 261, 353
primus, 132
prō, 55, 134
probāuit, 264
prōclius, 120
pūtāre, 383
quadruped-, 132
quam, 264
quattuor, 62

-que, 134, 239, 252, 261;
-que ... -que, 279
quī, 133
quīnque, 247
quis, 130, 134, 168; *quid*, 52, 129, 266
quisnam, 176
quisque, 134
quisquis, 130, 331
quod (conj.), 134
quom (Archaic), 134
regō, 279
relictus, 380
rēnouāre, 90
rēx, 60, 255, 322
Rōmae (loc.), 252, 256
rōs, 395
rota, 38
rubēfacere, 105
ruber, 121
rubēre, 91
rubēscere, 90
rūri (loc.), 252, 256
rursus, 133
sacrātur, 268
sal-, 60
salua seruassis, 262
saluus, 267, 398
sancire, 249
sarcire, 21
scabō, *scābi*, 308
sē, 130
secundus, 132
sedere, 55, 73, 249
sēdēs, 73
sella, 117
sēmen, 110
Sēmōnēs, 110
senectūt-, 112
seuor, 122
sepelire, 249
septem, 8, 248; *septimus*, 132
sequor, 251
seruire, 249
Seruius, 121
seruus, 114–15, 249, 254;
seruos, *seruom* (Archaic), 254
sex, 8; *sextus*, 132
similis, 131
simplex, 132
sine, 361
smus, 394
sistō, 90
sociae (dat.), 263
sōl, 111

sorbēre, 341, 393
 sorōr-, 60
 speciō, 70
 speculum, 257
 stabulum, 118
 status, 57, 248
 strātus, 248
 suānis, 253
 sub, 134
 sum, 87, 256; *esom*
 (Archaic), 87, 256; *est*, 58,
 60, 304; *erō*, 96; *erit*, 95,
 250; *esed* (Archaic), 250;
erunt, 95; *fuera*, 250;
fuert, 251; *siem* (Archaic),
 96; *siēs* (Archaic), 251,
 261; *sit*, 266; *sitis*, 251;
essēs, 251; *estō*, 404; *escit*
 (Archaic), 90; *esse*, 249,
 251, 253, 256; *futūrus*, 74,
 267
super, 134, 314
sūs, 356
sums, 130
tatod (Archaic), 261
ter, 133
terra, 52, 254, 255, 361
tertius, 132
testāmentum, 268
testimōnium, 118
testis, 22
Teutonī, *Teutonēs*, 300
-tiōn- (n. suff.), 347
toga, 70
tollere, 361
tonāre, 249
trāns, 289
triped-, 131
tū, 128
tūli, 250
tunc, 279
turris, 255
uādere, 348
uae, 135
-ue, 134
uehō, 386; *uehere*, 56; *uēxi*,
 257
ueiuont (Archaic), 256
uelle, 249
ueniō, 253, 360; *uenire*, 262
uentus, 64, 98, 195
uērus, 289
uia, 264
uicus, 228
uidua, 261
uieō, 375

uiginti, 356
uīnum, 38
uiri (gen.), 114, 247, 252;
uirum (gen.), 116
uirus, 61
uīsus, 261, 266
uītām, 264
uīuus, 407
umerus, 265
unde, 107
ūnus, 131, 254; *oinom*
 (Archaic), 254
uouēre, 264
uōx, 104, 195
ūrere, 254
ursa, 119
uxōrem dūcere, 18
Veneres (Archaic), 104

Faliscan

arcentelom, 263
carefo, 250
ceres, 263
-cne, 252
douiad, 263
efiles, 262
far, 263
ffiqod, 251, 263
hutulom, 262, 263
karai, 263
loufir, 263
Marci, 252
med, 263
melatom, 263
peparai, 250, 263
pipafo, 262
sociai, 263

Umbrian (and South Picene)

actud, 265
ander, 248
ape, 265
arsfertur, *afiertur*, 265
benust, 250
cerfer, 265
dede, 250
dirsa, 265
dupurs-, 123
emantur, 266
erom, 251, 265
esme, 267
esmen (S.Pic.), 129, 267, 385

faça, 265
fri, 267
frosotom, 266
fufans, 250; *fust*, 91, 251;
futu, 267
Grabouus, 266
heriei, 251
herter, 266
kumatir, 265
manf, 264
mers, 266
maneklu, 266
mutu, 264
nerf, 265, 267
nome, 267
ocrer, 266
peracrei, 267
peretom, 266
persei, 266
pesetom, 266
pihaciu, 267
pihafi, 97, 264
popler, 264
postin (S.Pic.), 267
prephohotatu, 266
prupehast, 250
prusikurent, 250
pustin, 267
safinūs (S.Pic.), 264
saluo seritu, 267
sei, 266
sihitu, 265
tefa, 265
testru, 265
tokam (S.Pic.), 267
toto, 264; *tota*, 267
Tursa, 264
ueiro pequo, 33, 267
ueres, 266
uerof, 116
uuro, 248
uirseto, 251, 266
une, 265
utur, *uin-*, 251
uze, 265
videtas (S.Pic.), 267
vitlaf, 265
vufru, 264
zeřef, 265

Oscan (and Sicel)

anamum, 248
aragetud, 268
biitam, 264
comenei, 114; *comono*, 115

dema-, 170
didest, 250; *deded*, 269
eisak, 269
éitiuvam, 268
fangvam, 248
fefacid, 250
fufens, 264
hipid, 250
húrz, 264
isídum, 269
kúmbened, 250
kúmbenniéis, 269
kvaisstur, 269
Mamertei, 105
meíai, 252, 264
Núvlanús, 115, 264
paam, 264, 268
patensíns, 264
-pe, 252
píbe (Sicel), 269
pítpit, 130
prúfatted, 264, 269
puf, 107
puklum, *puclom*, 214, 248
pútiad, 251
sakarakleís, 264
sakarater, 268; *sakrafir*, 264
stahínt, 264
tancinud, 269
toutad, 252
triúbúm, 269
trístaamentud, 268
úpsannam, 251, 269
viú, 264; *viái*, 252

Romance (French unmarked)

a (Sp.), 348
aveugle, 279
cantaré (Sp.), 258
cent, 258
cento (It.), 258
chanterai, 258
Clovis, 315
dent, 3–4
dente (It.), 3–4
dente (Port.), 3–4
diente (Sp.), 3–4
estar (Sp.), 361
kentu (Sardinian), 259
Louis, 315
oc (OFr.), 258
oil (OFr.), 258
pas, 6, 133, 331

pena (It.), 257
personne, 333
specchio (It.), 257
témom, 118

VI. Celtic

Common Celtic

**kamb-*, 279
 **karant-*, 276, 296
 **māros*, 276
 **yo-*, 130

Continental Celtic (Gaulish unmarked)

ambi-, 276
Anextlomārus, 35
Aremoricí, 121
arganto-, 39
Arto-, 276
bercunetacam (Celtib.), 279
cambion, 279
Carantomagus, 296
Caturix, 35
-cue (Celtib.), 278; *-cue* . . .
 -cue, 279
delgu, 276
dughontio, 130, 277
exsops, 123, 279
-io, 277, 279
iomni (Celtib.), 277
litano-, 276
-māros, 297
Mediolānum, 276
Moritasgus, 35
nelitom (Celtib.), 279
Nowodūnum, 122
oscuez (Celtib.), 279
Petrucori, 409
piissiumí, 91, 278, 279
ponc, 279
reguc, 279
-rix, 60
Segomāros, 276
silabur (Celtib.), 280
soisum (Celtib.), 129
sua (Celtib.), 279
tatuz (Celtib.), 95
Teutatis, 267
tidres, 120
tricanam (Celtib.), 279
toncsüontio, 279
trigaranus, 131
uer- (Celtib.), 279, 280
uertannei (Celtib.), 279
Uvamokozis (Lepontic), 35,
 123
Uxama (Celtib.), 247
uxello-, 290
Vercungetorix, 279

Insular Celtic

INSULAR CELTIC

**ro-*, 282
 **yo*, 148, 277

GOIDELIC (OLD IRISH UNMARKED)

a 'their', 281
ad-cíu, *ad-cear*, 82; *ad-cí*, 279
ad-fét, *ad-cuid*, *-éicdiá*, 287
ad-móingter, 290
ad-suidi, 90
ail, 130, 132
ainm(m), 111
aite, 19
an-, 56, 276
ara-rethi, 148
arathar, 118
arbor, 209
ard, 290
art, *Art*, 28, 276
as-beir, *-epur*, 288
as-boind, *-op*, 287
atdom-indnastar, 140
athair, 19, 275, 276
ben, 52, 275
berid, 284, 285, 286, 288;
 biru, 53, 275, 276; *-beir*,
 284, 286; *birt*, 281; *beraid*,
 247
bó, 284; *bom*, 60
bráu, 111
cach, 290, 296
camd, *cechan-*, 286; *cechnaid*,
 286
carae, 276, 288; *carait*, 296
cath, 284
cetheoir, 120
cia, 130
cín, 112
cír (Mid. Ir.), 52
cland, 290
cloth, 285
co, 134

cóc, 247, 280, 291
 com-arc, 294
 contherchomracc, 287
 cosmil, 281
 crenaid, 285; -cren-, -cria-, 359; -crith, 281
 cretim, 98; cretim, 296
 cride, 52, 276
 cú, 64
 Cunorix (Ogam), 35
 daig (Mid. Ir.), 275
 dán, 284
 daur, 52
 día, 22
 dingid, 285
 do-fúarat, -diúair, 287
 do-fuissim-, -tussimter, 290
 do-gní, do-gén, 286
 do-moinur, 55
 do-sluindi-, -diltai, 287
 druí, 109
 du-airngerat, 287
 ech, 284
 écsamli, 281
 eter, 134
 fadessin, 130
 fed, 290
 fén, 56
 fer, 285, 288, 291; fir (nom. pl.), 115, 285; firu, 115
 Fergus, 35
 -fess, 63, 275
 fí, 61
 fichid, fich, 93
 filed, 283
 finn, 291
 fír, 289
 flaith, 289
 flai(i)them, 289
 fo, 134, 297
 for, for-, 134, 294
 fo-tumndiriut, 287
 -géna, 91
 gnáth, 276
 Goidel, 283
 gonim, 53
 grán, 276
 guád, gígis, 286; gess, 281
 íasc, 290
 imb, 276
 In Dagdae oll-athur, 23
 is, 60, 289
 ith, 290
 lán, 98
 léicid, léicfid, 286
 loch (Sc. Gael.), 75

luchair, 55
 Lugudeccas (Ogam), 105, 283, 285
 luid, 92
 Luigdech, 283
 maicc (gen.), 283, 284
 maidid, 279
 mann, 290
 maq(q)i (Ogam), 114, 247, 277, 283, 284
 máthair, 19, 60
 melid, 163
 mes, 290
 mud, 53
 midiur, 322
 mil, 285
 mlicht (Archaic), 290
 moín, 290
 mór, 285, 290, 294, 297; már (Archaic), 276, 290
 móraid, 285; mórais, 281
 mumme, 19
 neb, 296
 ní, 55
 no-, 135
 óen, 131
 renaid, -ria, 286
 rethid, 287
 ro-, 55, 134, 282
 ro-chúala, 281
 salann, 60
 sé, 280
 secht, 280
 simiu, 122
 suir, 60
 sní, 296
 so-, 123
 sruth, 290
 suide, 285
 -tabar, 286, 287
 talam, 105
 tart, 276
 tégi, -téig, 285; tíagu, 53, 285
 teoir, 120
 tre, 289
 trichat-, 279
 tuáth, 284, 285, 288; tóth (Archaic), 285
 úachtar, 122
 úa hArt, 284
 úasal, Archaic ósal, 290
 velitas (Ogam), 283

BRITISH CELTIC

Cumobelinos, 281

WELSH (MODERN UNMARKED)

am, 276
 an-, 276
 ar (Mid. W), 282
 arth, 276
 ben, 275
 bryn, 291
 caffat (Mid. W), 281
 calaur (OW), 294
 cant (Mid. W), 281
 car, carho (Mid. W), 281
 Catmör (OW), 35
 cerant, 296
 ciglef (Mid. W), 281
 coueidid (OW), 294
 Cynfelyn, 281
 dant, 297
 diogel, 296
 duuw (OW), 296
 dyweddy, 294
 ercit (OW), 294
 fy, 127
 gell (Mid. W), 275
 go-, 297
 gochelyd, 297
 grawn, 276
 guir (OW), 291
 guor- (OW), 294
 gwae, 135
 gwastad, 57
 gwely, 297
 gwen (fem.), 291
 gur, 291
 gwyn, 291
 gwys, 275
 heno, henoeth, OW henoid, 294
 hynt, 296
 iau, 55
 iben (OW), 294
 issid (OW), 277
 kereis (Mid. W), 281
 llydan, 276
 maur, OW maur, 276, 290, 294, 297
 med (OW), 294
 meibion, 291
 mes, 290
 n, 296
 noeth, 294
 Peithuryr, 291
 plant, 290
 pump, 291

pwyl, 52
rhwyf, 296
ry-, 282
sauth, 292
sydd, 277
yd, 290
yn, 134

OTHER BRITTONIC (MIDDLE BRETON UNMARKED)

Breizh (Mod. Bret.), 294
brezhoneg (Mod. Bret.), 294
cret, 296
dans (Corn.), 297
diouguel, 296
Doe, 296
feynt (Corn.), 297
go- (Corn.), 297
goheles (Corn.), 297
guelas (Corn.), 297
ha (Corn.), 297
hent, 296
hins (Corn.), 296
me, 296
mur (Corn.), 297
nep, 296
ni, ny, 296
peb (Corn.), 296
pep, 296
pet, 296
querent, 296
roe, 296
ruuf (Corn.), 296
sur (Corn.), 297
tir, 296

VII. Germanic

Prehistoric and Inscriptional Germanic

Aflms (WGmc.), 106
**ga-* (n. pref.), 134, 322, 326
gl̥sum, 379
**gud̥-iga-*, 121, 318
harigasti, 309
**hlaibaz*, 300, 317, 368
**-iska*, 374
**-ipo-* (n. suff.), 319
**kuningaz*, 300, 368
**mik*, 135

**namnjan*, 90
**þik*, 135

Runic

ek, 310
gastiz, 303, 310, 329
haite, 313
hlewagastiz, 35, 310
-ka, 333
tawido, 310

Gothic

ada (Crimean), 313
amlif, 364
ans, 131
apukd̥iþs, 322
aiþis, 130
andbahtjan, 308
anþar, 317
armosta, 306
Attila, 370
augona, 315
ausona, 315
baifra, 89; *baifrans*, 97, 309;
baifran, 307
batista, 122
bidjan, 315
bitans, 98
broþar, 304
dags, 114–16, 305, 315; *dag*,
313; *dagis*, 129
faran, 308
Filippus, 6
fon, 111, 305
fra-, 134
frodoza, 306
fruma, 132
fullnan, 308
fulls, 98
ga- (v. pref.), 315
gadraihts, 322
ga-h-melida, 314
ga-laubjan, *ga-u-laubjats*, 313
ga-u-hwa-sehvi, 140, 313
giba, 119, 305
giban, 307
graban, 163
-h (conj.), 134, 314
habaida, 308
haiþait, 94
haiþs, 61
haiþreids, 65
haiþto, 314
haitan, 308
haris, 65
hausjan, 74, 315
himma, 364
hlaifs, 368; *hlaibans*, 314
huzd, 303
hwa, 313, 314
hwas, 130, 134
hwazuh, 134
ik, 127
im is ist, 87, 308
is (pron.), 130
jah, 314
malan, 164; *mol*, 93
man, 94
manags, *managiza*,
managista, 122, 306
mudjis, 121
miþ, 314
mizdo, 60
mops, 322
namnjan, 90
namo, 111
nauiþ, 314
nih, 314
niman, 314
qiman, 262
qipan, 314
sa so þata, 130, 306
saituan, 314
salbon, 308
satjan, 90
sawil, 111
siggwan, 302, 312
siggan, 312
silubr, 280
sitls, 117
skipa, 314
stainahs, 121
steigan, 53
swa, 279
tawjan, 310
þagkjan, *þahtedun*, 314
þamma, 129
þata, 314
Theodericus, 35
þis, 129
þu (paradigm), 128
þūsundi, 364
triggws, 313
twaddje, 313
twalf, 364
ufar-, 314
-uh, 314
unagands, 149
unte, 314
Vist-, *Vesi-*, 312

wahsan, 164
wai, 135
wairþan, 313
wait, 94, 314, 333
waljan, 308
warþ, *wairþun*, 306;
wairþum, 313
winds, 64

Old English

ācnāwian, 322
aelda (Northumbr.), 322
ærest, Northumbr. *ærist*, 322
alan, 322
āstelidæ (Northumbr.), 322
ā(w), 322
bana, 302
band, 330
barnum, 322
berende, 309
biddan, 302, 315
brōþor (gen.), 105
cēap, 318
cēosan, 319
cēowan, 317
cirice, 317
cūþ, 304
cwēn, 52
cyning, 300, 368
cynn, 322
dæg, 305, 315, 317
dic, 317
Drihten, *Dryctin*, 322
drincan, 330
eaht, 302
ēce, *ēci*, 322
efen, 330
Englaland, 316
eom, 308
fēond, 319
feorþa, 132
fif, 317
findan, 330
fisc, 317
fleax, 326
flēogan, 318
fōt, 319
Frea, 322
freosan, 317
gē, 306
gearn, 317
geldan, 317
geolu, 317
gesceap, 322

geþanc, Northumbr. *gidanc*, 322
gifu, 305
gīt, 306
gos, 317
guma, 22, 110
gūþwudu, 329
gydig, 121
hæfde, 318
hal, 322
hālig, Northumbr. *hāleg*, 322
hätte, 313
hehton, 308
helpan, 319
heord, 52
hēr, 304
hlāf, 314, 317
hlēo, 327
hlūd, 315
hridder, 118
hwæt, 52
hweohl, 117
hwī, 130
ic, 306
liofofa, *liofofost*, 306
lomb, 319; *lombru*, 320
mang, 317
Meotodes, Northumbr. *Metudæs*, 322
metian, 322
middangeard, Northumbr. *muddungeard*, 322
mōd, 322
mōdor, 60
moncynnæs (Northumbr.), 114
monn, 322
mūs, 61; *mȳs*, 326
ne, 55, 133
niht, 58
niman, 314; *numen*, 304
ofer, 317
ōr, 322
ōper, 317
ræt, 317
rīc, 322
rūn, 309
sacu, 330
sagu, 317
scieppan, 322
scip, 317
scyrte, 317
sē seo þæt, 306
sittan, 315
sōhte, 330
sōt, 73

stellan, 322
swefn, 118
sweostor, 326
þū, 306
Pumor, 330
Tig, 23
us, 317
wælisc, 292
wæs, *wæron*, 307
wæscan, 317
wāt, 94
wē, 306
wealas, 292
wearþ, *wurdon*, 307
weorc, *uerc*, 322
wild, 330
wit, 306
wlitan, 330
wurden, 307; *wurdon*, 313

Modern English

acknowledge, 322
(a)eon, 239
alone, 370
and, 134, 314
Andalusia, 312
Ararat, 337
barrn (Scots), 322
bane, 53, 302
bare, 74
barley, 263
basis, 229
be, 74, 267
bear 'carry', 53
beech, 40
believe, 313
bid, 302
bmd, 304
bitten, 98
blatherskite, 317
bone, 318
book, 324
both, 329
breadth, 5
brew, 404
bridegroom, 22, 110
bring, *brought*, 303, 308
broth, 404
brother, 304, 324, 353;
brethren, 305
burg, 340
but, 134
by, 107
-by (placename suff.), 328
call, 329

- change*, 279
cheap, 318
chew, 317
children, dial. *childer*, 319
church, 317
clay, 176
comb, 371
come, 262, 360
corn, 276, 380
cow, 341
crimson, 347
Cymbeline, 281
day, 317
deed, 304, 324
deep, 353
despot, 108
ditch, 317
dough, 263
dreg, 317
drench, 317, 319
drink, 319
Dutch, 300
early, 322
earth, 318
eenie meenie mney mo, 292
egg, 313, 317
eight, 304
-er (comparative suff.), 306
-est, 122, 306
fall, (to) *fell*, 319
fare, 325
farther, 122
father, 276, 304
fee, 19, 262
fiend, 309
filth, 5, 319
fire, 111, 305
first, 122, 132
fish, 290
flea, 395
flood, 324
foot, 51, 324; *feet*, 319
ford, 404
forlorn, 315
friend, 309
fright, 360
fro, 55
full, 304, 375
Gaul, 277
get, 317, 328
giddy, 121, 318
gift, 317
glass, 379
glossary, 229
guest, 317
haft, 302
Hagopian, 347
handwork, 322
hards, 52
he, 130
health, 319
hear, 74, 315
heart, 52
hepatitis, 108
here, 304
him, 326
hoard, 303
home, 318
hound, 341
hundred, 366
husband, 328
in, 134
-ing, 309
ingot, 53, 263
Iran, 187
is, 60, 304
-ish, 121, 374
-ize (suff.), 90
justice, 21
Khatchaturian, 347
knee, 52, 393
know, 355
Krishna, 189
lea, 383
learn, 318
lee, 327
light, 55
loaf, 300, 314, 368
long, *length*, 319
lorn, 315
loud, 315
lovelorn, 315
lox, 40
man, *men*, 319
marry, 317
may, 322
mead, 53, 294
Melissa, 229
mesh, 303
mud, 304
midwife, 314
Milan, 276
milk, 37
mind, 55
mood, 322
mouse, *mice*, 319
murder, 56, 304
nag, 317
needle, 118
nest, 73, 303
nether, 134
new, 383
nickname, 8
night, 58
nose, 75
now, 157
numb, 304, 314
numskull, 304
O'Hart, 284
of, *off*, 134
oh!, 135
old, 322; *elder*, *eldest*, 319
ombudsman, 239
one, 131, 370
only, 370
other, 386
over, 134
ox, *oxen*, 109, 305
Prakrit, 189
punch 'a drink', 131
queen, 52, 375, 386
quern, 38, 111
quoth, 314, 333
raw, 112
rune, 309
-s (pl. suff.), 5
sake, 330
salt, 60
salve, 393
same, 131
Sanskrit, 189
sap, 406
saw 'saying', 317
seat, 73
see, 314, 393
shaman, 352
shape, 322
ship, 314
shirt, 317
shit, 317
silt, 75
sing sang sung, 306
sister, 326, 328
sit, 55, 73; *sat*, 73
skiff, 314
skirt, 317
sloth, 5
snow, 228, 248
so, 279
son, 383
soot, 73
sow 'pig', 356
spew, 302
star, 302
stink, *stench*, 319
stoke, 72
stone, 318
stream, 117, 290

sun, 111
-t (2pl.), 320
take, 329
ten, 304
-th (noun suff.), 5
they, 329
thing, 324
think, 269, 314
thirst, 52, 182, 276
thorpe, 269; *-thorpe*, 328
tithe, 132
to, 239
tree, 52
Tuesday, 23
twig, 394
un-, 56, 133, 304
(un)couth, 304
under, 134
wagon, 56, 386
Wales, 292
warm, 117, 302
wasp, 74
water, 111, 305
wax, 182
Wednesday, 325
Welsh, 292
wet, 341
wheel, 36, 38
which, 130
white, 383
whore, 263
widow, 261
wife, 328
will, 303, 308
win, 189
wind 'breeze', 196, 354
wine, 38
woe, 135
wood, 290
work, 341
wot, 94
wrath, 332
wrench, 325
yard, 322
yarn, 317
yellow, 317
Yiddish, 325
yield, 317
yoke, 55, 228
youth, 118

Old High German (and Frankish)

ander, 317
Anselm, 35

bant, *buntun*, 306
beranti, 309
bibar, 112
birenkit, 325
däutisc, 121
ende, 325
farhali, 117
fibu, 19
fimf, 317
flahs, 326
fliogan, 318
Fridugundis, 35
Früa, 326
geba, 305
gilouban, 313
habēm, 308
Hadubrant, 35
bēr, 304
Hiltibrant, 35
hintagu, 130
**Hlodowig* (Frank.), 315
Hludwig, 35
hlūt, *hlūd*, 315
im, 326
Liutpold, 35
nemmen, 90
rotēn, 91
selb selbo, 407
smr, 60
snuster, 326
swāgur, 117
tag, 305
umbi, 239
uns, 317
untar, 134
uorun, 325
Waldram, 35
Waltbert, 35
wilt, 303
zweio, 313

Middle and Modern High German (Modern unmarked)

Aar, 167
acht, 304
adelar (Mid. HG), 167
Adler, 167
auch, 333
bauen, 386
Bein, 318
Berg, 279
Bruder, 324
Buch, 324
Deutsch, 121

dich, 135
Ding, 324
ewig, 322
Flut, 324
forschen, 90
fragen, 214
Frau, 322
fuhren, 325
Fuß, 324
Gans, 317
Garn, 317
gelb, 317
gelten, 317
gerenkt, 325
gewiss, 63
glauben, 313
Gottfried, 35
Hase, *Hasen*, 305
Heim, 318
heiße, 313
Herz, *Herzen*, 314
heute, 130
Hirse, 112
ich, 407
Jahr, 330
jung, 330
kauen, 317
Kauf, 318
Kirche, 317
Laub, 317
laut, 315
lieber, 326
legen, 354
nach, 135, 407
mut, 107, 314
Mut, 322
nehmen, 314; *nahm*, 333
nennen, 90
noch, 314
raten, 333
Ratte, 318
Roggen, 404
Saft, 406
schaffen, 322
schwer, 172
Schwester, 326
sehen, 314
sind, 304
singen sang gesungen, 306
Stem, 318
stellen, 322
Sülze, 75
Tag, 315
Tat, 304, 324
tränken, 317
um, 239

und, 314
 Vieh, 262
 wahr, 289
 Wald, 327
 weiß, 333
 wurden, 330
 Ziege, 393

Other West Germanic (Old Saxon unmarked)

cuning, 300
 fif, 317
 fiskar (E. Frisian), 115
 flas, 326
 gumono, 327
 hlêo, 327
 liobora, 326
 manno, 327
 mûs, 326
 uuald, 327
 uuros, 327
 wagen (Mid. Dutch), 56

North Germanic (Old Norse unmarked)

-a, -at (neg. suffixes.), 331
 ár, 330
 áss, æsir, 330, 333
 batt, 330
 berande, 309
 berr, 331
 býr, 328
 dáð, 304
 dagr, 305, 315
 díki, 317
 drekka, 330
 egg, 313
 ekki, 331
 em est es, 308; er, 308
 engi, 331, 333
 faðir, feðr, fœður, 105, 331
 finna, 330
 fljúga, flýgr, 330
 framr, fremri, 330
 Frigg, 326
 fullr, fylla, 330
 gater (Norw.), 332
 gestr, 310, 329
 -gi, 331
 gøtur, 332
 greip, gripo, 306

hánn, hann, 135
 heiti, 313
 hér, 304
 heyra, 315
 hvat, 330
 hvergi, 331
 -(i)nn, -(i)n, -(i)t, 331
 jafn, 330
 jöunn, 330
 -k, 333
 kallask, 331; kǫllumk, 331
 koma, kóm, 330, 331, 332
 kommer (Norw.), 332
 kvað, 333
 líta, 330
 lyngfisk, 329
 mik, 331; -mk, 331
 mjök, 133
 nam, 333
 ok, 333
 orð, 330
 óss, 322
 ráða, 333
 Ráðulfr, 35
 róa, roru, 330; rera, 308
 rún, 309
 saga, 329
 sík, -sk, 331
 skáld, 329
 skor, 305
 skyrta, 317
 slyngva, 330
 spök, 330
 sótti, 330
 staðr, 57
 þorp, 328
 Þorr, 330
 þreifask, 333
 tryggr, 313
 tveggja, 313
 Týr, 23
 ungr, 330
 urðu, 330
 vatn, 111, 305, 330
 vent, 333
 villr, 330
 vóro, 330
 vreiðr, 332

VIII. Armenian

acem, 342, 345
 aganum, 342
 akn, 348
 al, 60, 342

am, 343
 anasoun, 98
 and, 344
 angorc, 341
 anicanem, 345
 anoun, 111, 402
 -arar (n. suff.), 347
 araur, 342
 arbi, 341; arbil, 342
 areg, arev, 348
 argeloum (paradigm), 345;
 argelić, 346
 as, 344
 asem, 98; asē, 348
 ast, 344
 awadik, awanik, awasik, 344
 awl, 342
 -awor (n. suff.), 341, 347
 awr, 342
 ayc, 342
 ayd, ays, ayn, 344
 aydr, 344
 ayr, 57, 71, 342, 402
 bam, 345
 ban, 343
 banber, 123
 barjr, 340
 berem, 53, 341, 345, 346;
 berim, 346; eber, 91, 343,
 346
 cer, 342
 cin, 340
 cnem, 349
 cov, 176, 347
 coumr, 52
 -d (demonstr. suff.), 344
 da, 344
 dast, 338
 dēz, 340
 dnem, 341
 dow, 128
 dowstr, 342
 doyn, 344
 dpanem, 340
 -ean (patronymic suff.), 347
 ei, 345
 end, 134, 348
 anker, 347
 erek, 342; erekoy, 348
 erek', 341, 342
 erkar, 342
 erkn, 342, 347
 erknčim, 342
 erkow, 341
 es, 127, 349
 ew, 134, 341, 347, 348

- ewt'n*, 341
gam, *gay*, 348
ger i veroy, 341
gereal, 98
get, 341, 343; *getoy*, 114;
 gets, 342
gir, 343
gitem, 341; *egit*, 92
glem, 345
gorc, 322, 341
gorcem, *gorcec'i*, 346
grem, 343, 345; *greal*, 346
hanem, 342; *hani*, *ehan*, 346
harbil, 342
harc'anem, 345
haw, 342
Hay, 337
hayec'ay, 346
hayr, 342; *haur*, 344
hecanim, 55
hin, 342
hing, 341
hmoc', 111, 344
hogi, 342
hoz, 342
hotn, 51
hour, 111, 341, 344, 348
i, 134
imanam, 55
inj, 127
i ver, 134
jerbakal, 344
jern, 117, 341
jern, 112, 338, 344
jnem, 53
jnjem, 341, 345
jowkn, 61, 342
joyl, 53
kalni, 341
k'ani, 52
karmir, 347
kin, 52
kou, 60, 341
k'own, 118; *k'noy*, 343
k'oyr, 341, 342
k'san, 342
lizem, 341; *elez*, 343
lk'anem, 345; *eluk'*, 340, 341,
 342, 346
lowsawor, 343, 347
loys, 55, 342, 343
mard, 56, 349
matean, 343
matenagir, 343
mauwrows, 348
mayr, 60
mæg, 341, 343
me'anim, 345
mez, 340
mi, 131, 342, 348, 361
mugamac, 343
-n (demonstr. suff.), 344
na, 344
ni, 134
noym, 344
nstay (paradigm), 346
nu, 60
ogi, 342
oßj, 341
oror, 342
orsam, 345; *orsac'i*, 346
ost, 342
otk', 342
otn, 341, 344
ownër, 347
owtem, 345, 348
-owt'ium (n. suff.), 347
p'etowr, 341
p'ok'r, *p'ok'ownk'*, 344
-s (demonstr. suff.), 344
sirt, 52
skowmd, 341
soym, 344
srtiw, 107
sterj, 341
tam, 341, 349
tan, 108; *towm*, 108
t'aranim, 52
tasn, 340, 341
telowoj, 107
tëz, 343
tkar, 347
towrk', 342
tram, 52
xalam, 337
xnjor, 337
z- (dir. obj. pref.), 347
zgenowm, 345
zgest, 343
zmanne, 340
kaßi, 354
kaß, 354
klyosam, 359
klyu, 354, 360
knä-, 355
ko, 353
ku-, 354
kukäl, 117, 354
k'älli, 357
kumseñc, 355, 360
kus, 354
lac, 92, 354
lake, 354
lu, 362
luk-, 55, 353
lyäk, 93; *lkätsi*, 97
mā, 360
musnätär, 358
nam, 360
nämseñc, 358
näş, *ñuk*, 358
nas-, *naş*, 360-1; *nestem*,
 85-6
ñom, 111, 360
ñom-klyu, 34, 360
ñom tä-, 34
nu, 135
ñu, 354
ok-, 353, 356
or, 353
păcar, 354, 355
pär-, 53
pärkär, 354
pe, 51
pkänt, 356
pontsäm, 355
poto, 360
pracar, 353, 355
prak-, 353
praski, 360
p'kkäl, 361
ritwāl, 98
rtär, 353, 355, 356
šäk, 360
säle, 60
šäm, 355
şaptänt, 132
sas, 131
šašärs, 93
satkatär, 360
şpal, 354
swär, 353
swiñc, 358
tämyo, 361
tkam, 354
tpär, 353

IX. Tocharian

Tocharian A

- ak*, 353, 355; *ašäm*, 357
ak-, 353, 355
ekär, 356
känt, 353, 354
käntu, 354
kanwem, 52

trāṅkmār, 85–6
 tsāk-, 354
 want, 354
 wäs, 355
 wiki, 356
 wše, 356, 360
 yakwe, 354
 yārḱ, 360
 yaš, ymām, 83
 yāt-, 360
 ymār, 361
 yuk, 354

Tocharian B

aikare, 356
 aise, 356
 aiskau, 358
 -aiwenta, 357
 āk-, 353, 355; *akem*, 358
 āke, *akenta*, 356
 alyek, 130
 auk-, 353, 356
 camel, *cmela*, 356
 ek, 353, 355; eš, 361; ešane, 357; ešnesa, 361
 ikām, 356
 kacāp, 361
 kālywe, 360
 kante, *kante*, 353, 354, 361, 366; *kāntenma*, 356
 kantwo, 354
 kārnāstār, *kārnātār*, 359
 kārsanam, *šarsāššam*, 358
 kāryāñ, 52
 kautāmai, 85–6
 keṃ, 354; *kenā*, 361
 kešcye, 354
 kest, 354
 ke_u, 353
 kewu, 53
 klyausau, 359; *klyaušim*, 358
 knā-, 355
 kokale, 354
 ku-, 354
 k_ulse, 52, 354
 lareñ, 358
 lareññentār, 358
 leki, 354
 lipetār, 358
 lkātsi, 97, 356
 ltu, 98
 luk-, 55, 353
 luwo, 362; *luwāsa*, 355
 lyautai, 361
 ma, 361

mācer, 354
 mit, 53
 musnātār, 358
 nās, 127
 ñem, 360
 ñu, 354
 onohme, 362
 or, 353
 osotār, 358
 pacer, 354, 355, 357
 pällent, 354
 pār-, 53
 pārwešše, 354
 pik_ul, 361
 pinkem, 358
 prek-, 353; *preksa*, 92
 procer, 353, 355
 proskiye, 360
 ratre, 355, 356
 rmer, *ramer*, 361
 riār, 353
 šak, 360
 šamāne, 352
 šana, 52, 355
 sātketār, 360
 šäyeyca, 361
 še, 131
 se sū te, 130
 snai, 361
 stare, 87
 suwo, 356
 swäre, 353
 tākoy, *takoy*, 361
 tallam, *tālašši*, 361; *talläššām*, 358
 tapre, 353
 tsāk-, 353
 tsāk-, 354
 tsälpetār, *tsälpäššām*, 358
 tuwe, 128
 wase, 355
 utentse, 356
 yakwi, 115
 yapoy, *ypauna*, 356
 yarke, 360
 yente, 354
 ysamna, 362
 yšiye, 356, 360

X. Balto-Slavic

Common Slavic

*küñezī, 300
 *otū, 114
 *xlēbū, 300

Old Church Slavonic

azū (paradigm), 127
 bēaše, 375
 berp, 53, 89, 365, 372, 376;
 berpšt-, 97
 bilū, 346
 bljudp, 371
 bogatū, 121
 bogū, 121, 368
 bolji, 52; *bolš-*, 122
 byti, 375; *by*, *bystū*, 375
 celiū, 322
 cēna, 370
 češo, 168
 četvritū, 132
 čko, 130, 370; *česo*, 129
 dati, 97; *damī*, 341, 372
 desęte, 56, 371, 385; *desęte*, 112
 dēti, 371
 devęti, 385
 dinī, 375
 dīnīnica, 371
 dobrodějanje, 122
 dobrū, *dobruji*, 367
 drēvo, 52
 emu, 375
 glagolū, 117, 373
 glava, 119, 372
 gnęzdo, 55
 gosī, 371
 goveždi, 60
 gradū, 371; *grada* (gen.), 114;
 grada (nom. du.), 115
 ide, 374
 ime, 64
 mū, 131
 -isk- (n. and adj. suff.), 374
 jastū, 63
 jesmī, 87; *jestū*, 60, 375
 junū, 370
 kamy, 365
 kosa, 52
 küde, 107
 küñezī, 368
 küto, 52, 130
 lēvū, 371
 lomiti, 386
 luča, 55
 mater-, 60
 medū, 53
 medvędi, 123
 mīglo, 371

mínjo, 55
míza, 60
mrakú, 372
myši, 371
ne, 55, 133
nebes-, 169
nesti, 90; *neslú*, *nesení*,
něsomú, 98
niziú, 134
noč, 366
nositiú, 90
nosú, 370
nú, 135
oko, 315; *oči*, 105
onú, 130
osmi, 373
osmú, 132, 386
otiči, 370
otú, 374
ovíca, 370
peť, 373, 376
plünü, 375
pogreb, 163
polož, 375
povitiú, 375
prívu, 132
pro-, 55
rabú, 114–16
réxú, 370
ročka, 370
roka, 375
rúďeti, 91
rúmiškiú, 121
rútiú, 371
rúžda, 370
se, 130, 375
sěďeti, 55
sedmú, 132, 386
sěno, 383
sí, 130
slava, 368
slovesa, 106
služiti, 370
snúxa, 60
solí, 60
spštejo, 375
srídice, 52, 365, 370
stati, 371; *stanjo*, 372
stignoti, 53; *stignjo*, 371
sú, *sú-*, 123, 375, 383
súpa-, 250
súto, 368
svetiú, 371, 386
svítiti se, 383
svobodí, 400
svojí, 130; *svoj*, 375

synú, 375; *synu*, 371;
synúma, 107; *synúmi*, 106;
synúmi, 107; *synúmiú*, 106
tuři, 371
ty, 128, 386
tysšiti, 364
umirati, 371
usúnoťi, *usúpe*, 370
uxo, 370
varú, 302
věglasú, 123
vesti, 370; *vedjo*, 366;
věsomú, 92
vezti, 386; *vezjo*, 56
viši, 383
vlikiú, 56
vrixiú, 380
vü(m-), 134
vixodiú, 371
že, 135, 374
žena, 52, 119, 370; *ženojo*,
 375
ženo, 53
znati, 368, 371
zobú, 371

Russian

bern, 376
bog, 368
Boleslavú (ORuss.), 35
buzina, 40
byšěti (ORuss.), 91, 372
čitaju, *pročitaju*, 372
čtenie, 371
demica, 371
desjat', 371
golova, *glava*, 371
gorod, *-gorod*, 365, 366, 371
-grad, 371
gus', 365, 371
kakoj, 290
knjaz', 300
mglo, 371
mirtvú (ORuss.), 56
moroka, *mrak*, 371
ot, 114
Perumú (ORuss.), 23
písal, 62
pisál, 62
pjat', 376
rot, *rta*, 371
russkij, 121
serebro, 280
seryj, 370
slušať, 365

sneg, 228, 365
sobaka, 368
spat', 370
svjatoj, 371
tomu, 385
usnut', 370
vernus', 372
vino, 38
vosem', 370
vzghad, 371
zub, 371

Other Slavic

Bohuslav (Cz.), 35, 123
brěza (S.-Cr.), 366
dziesięć (Pol.), 371
ezik, *ezikot* (Mac.), 378
geś (Pol.), 371
grabowy (Pol.), 266
kniga, *knigata* (Mac.), 378
osiem (Pol.), 370
podzim (Cz.), 383
pün (S.-Cr.), 367
šerý (Cz.), 370
švęty (Pol.), 371
vreme, *vremeto* (Mac.), 378
vúk (S.-Cr.), 367
zab (Pol.), 371
zět (S.-Cr.), 366
złoty (Pol.), 365
zúb (S.-Cr.), 366
žut (S.-Cr.), 367

Baltic

LITHUANIAN

akis, 380; *ákys*, 384
akmuõ, 365; *asmuõ* (dial.),
 365
af, 135
arti, 383
ąš, 127, 380, 384
ašis, 380
asmas, 386
ašva (OLith.), 380
augusi, 98
aúkas, 383
aušrà, 380, 383
bėbrus, 112
bėgtu, 380
bėrzas, 366
būti, 97, 380, 384; *būs*, 91
debesis, 169
dervù, 52

dēšimt, 56, 384, 385, 386
 devyni, 385; deviņtas, 132
 diēvas, 22
 dubūs, 113, 365
 dukterēs (dial.), 105
 duona, 119
 dūoti, 380; dūostant-, 91
 dvjlika, 364
 ēsti, 60, 381
 ēsti, 63
 gēltas, 365
 genū, 53
 gēras, gerāsīs, 367
 gerū, 365
 gēve, 384
 giñti, 366
 gīvās, 380
 isaušti, 383
 jāntis, 383
 jūngas, 55
 jūngti, 366
 kāna, 370
 kās, 52, 130, 365
 kasā, 52
 kitas, 386
 klausyti, 365, 380
 liēpti, 386
 likti, 380; lieki, 366
 lizdas, 55
 lōju, 89
 māt, 128
 mēdis, 381–2
 medūs, 53, 366, 380
 menū, 55
 -m(i) (OLith.), 128
 mirtis, 56
 mótē, 60, 380, 381
 mótina, 386
 naktiñ, 107
 naūjas, 383
 nē, 55, 133
 nēšamas, 98
 nu-jam-lūdusam (OLith.), 148
 paršēlis, 117
 pavāsarīs, 383
 pēkus, 386
 Perkūnas, 23
 piāuti, 383
 pilnas, 98, 367
 pirmas, 132, 386
 prā, 55
 rankā, 381, 384
 rýtas, 384
 salia (dial.), 60
 saūsas, 61

šēmaīs, 383
 semiū, 290
 septuonas, 386
 sésti, 365
 šēnas, 383
 šimtas, 366
 širdis, 52, 365, 366, 380
 šis, 130, 364
 smākras, 348
 stergiuos, 53
 siū, 375, 383
 sūka, 381
 sūnūs, 381–2, 383; sūnumis, (OLith.) sūnumus, 106
 šuō, 104, 106, 380–1
 svarūs, 172
 šveñtas, 386
 šviesais, 383
 švytioti, 383
 tamē, 129
 tāmma, 385
 tėvas, 386
 tū, 128, 386
 tūkstantis, 364
 turēti, 386
 ugnis, 195
 ūšes (dial.), 386
 vedū, 366
 vežū, 56
 vienuolika, 364
 vilkas, 56, 366, 367
 viršūs, 380
 visas, 380, 383
 výras, 114–16, 381–2
 žaības, 366
 žasīs, 365
 žēltas, 365, 367
 žēntas, 366
 žiemā, 384
 žirnis, 380
 žmogūs, žmōnes, 383
 žmuō (OLith.), 383; žmune (OLith.), 105

LATVIAN

acs, 380; acis, 384
 ass, 380
 bēgt, 380
 būt, 380, 383
 ceļš, 384
 desmit, 384
 duot, 380
 dzērve, 384
 dzivs, 380
 es, 384

klausīt, 365
 likt, 380
 māte, 380
 medus, 380
 mīlme, 404
 par, 385
 rīs, 383
 ruoka, 384
 -s (refl. particle), 384
 septītais, 386
 sirds, 380
 suns, 380
 tam, 385
 tūvs, 386
 vai, 135
 ziema, 384
 zirnis, 380

OLD PRUSSIAN

ackis, 380
 anters, 386
 asmus, 386
 assis, 380
 buton, 380
 butten, 386
 datwei, 380
 deiwas, 114; deiwans, 116, 386
 dessimpts, 386
 en, 134
 genna, 52; gennan, 386
 gi(j)wans, geywans, 380
 gwāntei, 105
 gorme, 117
 grecon, 116
 ka-, 387
 kittans, 386
 luntwey, 386
 meddo, 380
 newints, newyntz, 385
 pallapsaey, 386
 Pats kailst, 322
 pecku, 386
 per, 134
 pirmois, 386
 poklausimanas, 98, 385
 pra, pro, 134
 salobisquan, 386
 schismu, 129
 septmas, 386
 seyr, 380
 smu(n), 383
 stesmu, 385
 sunis, 380
 swyntintwey, 386

syrne, 380
tauwyschen, 386
thawan, 386
tou, 386
tur, 386
usts, *wuschts*, 386
westwey, 386

XI. Albanian (Tosk unmarked)

Arbëresh, 390
athët, 393
besë, 20, 393, 394
bie(r), 394; *bjer*, 53
bredh, *brodha*, 396
dha, 393, 396; *dhashë*, 396;
dham, 394
dhe (Old Geg), 397
dhëmb, 393
dhetë (Geg), 394
dj, 393
dimër, 393, 395
djeg, *dogia*, *digjem*, 393,
 396
drekë, 394
drithë, 394
dru, 52
dubem, 396
ëmbi (Old Geg), 397
emër, 394
ferr, 394
fluturón (Old Geg), 398
gardh, 393, 394
gjallë (Old Geg), 398
gjalpë, 393
garpër, 393
gjashhtë, 393
gjerb, 393
gji, 394
gju, 393; *gjiu* (Geg), 395
gjumë, 393, 394
gjymtyrë, 394
be(l)q, 393
hie, 393
jam, 394, 396; *ështëë*, 393,
 398; *është* (Geg), Old Geg
anshtë, 398; *jeshë*, 396-
juve (Old Geg), 397
ka, *qe*, 393
kam, 393, 394, 396, 397
kë, 52, 393
kle (Old Geg), 396
lidh, 394

ligë, 393
llanë (Geg), 394
Llezhder, 394
mal, 395
marr, *merr*, 396
mbë-, *mb*-, 393
mbesë, 394
mbijell, *mbolla*, 396
mbledh, 393, 394; *mblodha*,
 93, 396
mbret, 394
mu, 394
mucl, 394
motër, 60, 394
murgu, *Geg mungu*, 392
na, 394
natë, 393, 394
ndër, 134
ndez, 393, 394
ne, 395
ngjesh, 394
ngrenë, 393, 394
ngroh, 393, 394
nip, 394
njeri, 394
nuse, 60
paça, 396
pemë (Old Geg), 397
për, 134
përposh, 51
pesë, 393, 394
pi, 394
piek, *poqa*, *piqem*, 393,
 396
pjerdh, 394
pljesht, 395
plotë, 394
presh, 393
që, Old Geg *qi*, 397
qiell (Old Geg), 398
rashë, 396
rë (Geg), Old Geg *ren*, 405
resh, 395
rrunjë, 394
sbi, 393
shkallë, 393
shoh, 393
shpirt (Old Geg), 398
shtatë, 393
shtazë, *Geg shtasë*, Old Geg
shtanse, 392
shteg, 53, 393, 395; *shtigje*,
 393
sy, 393
ter, 52, 393
tha (Old Geg), 397

thanj, 395
them, 393, 396
thi, 393, 394
uje, *ujera*, 395
ujk, Old Alb. *ulk*, 393
vë, *vura*, 396
vera, *Geg vena*, 392
vesh, 394
vjedh, 56
ynë (Old Geg), 397
zjarm, 117, 393
Zoti (Old Geg), 397

XII. Unaffiliated

Phrygian

abberet, 53, 402
addaket, 402, 403;
addaketor, 402
akkalos, 403
akke, 403
anar, 57, 71, 402
autos, 404
bekos, 401
edaes, 91, 403
egedou, 403
egeseti, 402
eitou, 403, 404
eneparkes, 403
gegaritmenos, 404
gegremeinan, 403
glouros, 402
ios, 403
kakoun, 403
ke, 402
mater, 402
oi, 403
onoma, 402
podas, 402
protuss[e]stamenan, 403
Sabazios, 400
semoun, 403
tetikmenos, 403
tios, 403
zemel-, 402

Thracian

brza, 404
br(o)utos, 404
Diazenis, 400
Dizazenis, 400
-para, 404
Sabazios, 400

Macedonian

abroutes, 405
Ber(e)nika, 405
Bilippos, 405
danon, 405
dōrax, 405

Illyrian

Deipaturos, 406
grabion, 266
rhinos, 405
sabaia, *sabatium*, 406
teuta-, 406
Vescleves, 35, 406

Venetic

alkomno, 407
andeticobos, 406
atisteit, 406
deivos, 407
donasto, *donasan*, 406
ego, 407
ersiniot, 407
horvionte, 406

-kve, 407
louzerophos, 406
mego, 407
murtuvoi, 407
sselboisselboi, 407
rolar, *toler*, 406
vhagsto, 406
Vineti, 114
vivoi, 407

Messapic

damatras, 407
dazimaiti, 114, 408
hipades, 408
hipakaθi, 408
kl(a)ohi, 408
logetibas, 408
pido, 408
-θi, *-si*, 408
Zis, 408

Lusitanian

doenti, 409
porcom, 409

taurom, 409
Veaniticorī, 409

XIII. Non-Indo-European

(Akk. = Akkadian; Ass. = Assyrian; Bab. = Babylonian; Finn. = Finnish; Geor. = Georgian; Heb. = Hebrew; Hurr. = Hurrian; Phoen. = Phoenician; Sem. = Semitic)

baprunnu (Bab.), 184
barittanu (Bab.), 184
binkarannu (Bab.), 184
ghvmo (Geor.), 38
hinzuri (Hurr.), 337
išḫu(l)lu (Ass.), 158
išpa(t)ta(l)lu (Ass.), 158
karmil (Heb.), 347
kuungas (Finn.), 300
pilakku(m), *pilaqu(m)* (Akk.), 73
**qadm* (Phoen.), 225
**wayn-* (Sem.), 38

Subject Index

- ablative, 104–5, 106–7, 114, 116
ablaut, 73–4; origin of, 74
absolute constructions, 149
absolute verbs (Celtic), 282, 286–8
accent-ablaut classes, 107–10
accent and stress, 62; *see also* mobile accent system
accent retraction, 104, 114; in deriving nouns from adjs., 120
accented zero-grades, 74
accusative, 103, 104, 105, 106, 114, 115–16
Achaemenid dynasty, 210, 337
acrostatic, 107–8
acrostatic presents *see* Narten presents
active, 82
adjectives: comparison of, 121–2; formation of, 120–1; pronominal, 130
admirative, 395–6
adverbs, 139; formation of, 132–3
Aeolic, 223
Aequian, 269
Aesop, 186
Afrikaans, 328
afterlife, 25, 41
Agamemnon, 27
agent-nouns, 111–12
Agni, 24
Ahura Mazdā, 205
Akkadian, 212
Akkadograms, 160
Alaksandus, 168
Alanic, 217
Albanian: law, 392; morphology, 395–7; phonology, 392–5; prehistory of, 390–1; syntax, 397; *see also* Geg; Old Albanian; Tosk
Alexander the Great, 216, 235, 404
Alfred the Great, 316
allative *see* directive
alliteration, 33, 151, 176, 261, 310
alphabets, 13
amphikinetic, 107–8, 109, 208
analogy, 5
Anatolia, 337; as location of IE homeland, 43
Anatolian: categories absent from, 155, 156; clitic subject pronouns in, 129; establishment of, 154; gender, 102–3; morphology, 156–7; *nt*-participle, 97; phonology, 156; position of in IE, 155; syntax, 157–8
Andronovo culture, 42
Angles, 316
Anglo-Frisian, 317
Anglo-Norman, 258
Anglo-Saxons, 292, 294, 316
animals: domestic, 37; wild, 39
animate, 103; Slavic, 372–3
Anittas, 158
Anquetil du Perron, A. H., 204
Anthony, David, 41–2
aorist: absence from Anatolian, 155; stems, 92–3
Apam Napāt, 24
Aphrodite, 22
apophony *see* ablaut
Appendix Probi, 257
ā-preterite, 250
Aramaic, 215, 324; alphabet, 215
Arcado-Cypriot, 222, 224, 232
archaeology, Indo-European, 35–44
archaic language: in law, 21; in religion, 26
Archaic Latin, 252
archetype, 7
Armenian: alphabet, 338; deictics, 344; nouns, 343–4; phonology, 340–3; verbs, 344–6; *see also* Middle Armenian; Modern Armenian
Armenian Consonant Shift, 340–1
Armenians, 401
Arsacid dynasty, 215, 337
ārya-, 19
Aryan, 19; history of usage of term, 187

- Arzawa letters, 154
 Aśoka, 196
 aspect, 83; Slavic, 371
 Assamese, 199
 Assyrians, 158
ā-stems, 119; in Sabellic, 264
ā-subjunctive, 276, 286, 358–9; in Celtic, 251;
 in Italic, 251
āśvamedha, 24
 Ásvins, 23
 Átar, 24
 athematic inflection, 77–8, 87; nouns,
 103–10; presents, 87–8; replacement by
 thematic inflection, 78
 Attic, 223–4
 Attic-Ionic, 224
 augment, 91–2, 231, 346, 403
 Avesta, 204–5, 216
 Avestan: alphabet and orthography, 206;
 differences between Old and Younger, 204,
 208; evidence for mobile accent system,
 207–8; manuscript tradition, 205–6;
 morphology, 208; phonology, 206–8; *see*
 also Old Avestan; Younger Avestan
 Bactrian, 216
 bahuvrihis, 122–3, 194
 Balkan speech area (*Sprachbund*), 378, 392,
 395
 Balochi, 217
 Baltic, 378–81; morphology, 380–1;
 phonology, 380; *see also* Balts
 Balto-Slavic: morphology, 367–8; numerals,
 364, 385; phonology, 365–7; subgrouping
 of, 364
 Balts, 378–9
 Bartholomae's Law, 118, 181, 188, 204
 "basic word order", 141
 beech, 40
 Behaghel, Otto, 34; Behaghel's Law, 34
 Behistun *see* Bisotun
 Belarusian, 376
 Bellerophon, 27
 Bengali, 199
 Benveniste, Émile, 180
Beowulf, 151, 316
 *-bh- (case element), 106–7, 156, 364
 Bihari, 199
 Bisotun inscription, 212, 213, 337
 Boğazkoy, 154
 "Book Pahlavi", 206, 215, 216
 Bopp, Franz, 280–1, 301
 borrowing(s), 1, 6, 72–3, 158, 168, 199, 212,
 216, 217, 236, 260, 295, 300, 320, 325,
 328–9, 332, 337, 338, 339–40, 352, 365,
 368, 371–2, 376, 378, 390
 Bosnian, 377
 Botai culture, 41
boukólos rule, 64, 133
 Brāhmī alphabet, 197–8, 216, 353, 354, 355
 branching noun phrases, 151
 breaking, 318, 329–30
 Breton, 275, 294–6; morphology, 295–6;
 syntax, 295
 bridewealth (bride price), 18
 "brightening", 317–18
 British Celts, 292
 Brittonic: phonology, 291–2
 Brown, Dorcas, 41
 Brugmann, Karl, 183; Brugmann's Law, 183
 Buddhism, 196–7, 215, 351
 Bulgarian, 378
 Burgundian, 312
 burial practices, 25, 40–1; Andronovo and
 Indo-Iranian, 42
 Buzuku, Gjon, 391, 397
 Caesar, Julius, 18, 274, 277
 caesura, 31
 Caland, Willem, 123; Caland system, 123;
 "Caland's Law", 123
 "canonical creature", 326
 Carian, 148, 176
 case(s), 102; development of additional, 102,
 381; functions of, 102; *see also under*
 names of specific cases
 case endings: athematic, 103–7; thematic,
 113–16
 case-like elements, 107
 caste system (India), 17–18, 27, 195
 Castilian, 259
 Castor and Pollux, 23
 Catalan, 259
 Caucasian languages, 217, 339
 causative-iteratives, 90; in Italic, 249
 Celtiberian, 277–8
 Celtic: morphology, 276; phonology, 276;
 relative pronoun in, 148; syntax, 277;
 verb-initial order in, 142, 144
 Celts, 352; prehistory of, 274; society, 18
 centum, 53–4
 chains of clitics *see* clitic chain(s)
 change, linguistic, 4–5; misconceptions
 surrounding, 5; sources of, 4–5; types of, 5
 characterized presents, 88
 Charon, 25
 Chimaera, 27
 Chinese Turkestan, 215, 216, 351
 Chomsky, Noam, 13, 137
 Choresmian, 217
 Christianity, 258, 274, 283
 Cisalpine Celtic *see* Lepontic

- classes of PIE society, 17
 Classical Armenian, 338; *see also* Armenian
 clause(s), syntax of, 140–1, 147–8, 165–6
 clitic(s), 62, 99, 129, 135, 146–7, 208, 313,
 331; positioning of, 235; syntactic behavior
 of, 140
 clitic chain(s), 99, 147, 165, 231, 313;
 Anatolian, 157
 clitic doubling, 174
 clitic groups, 149–50
 clitic raising, 239
 clitic subject pronouns, 156
 cliticization, 150
 Clytemnestra, 27
 Code of Manu, 186
 cognate, 4
 collective(s), 102, 118, 143; collective pl., 106,
 156
 coloring, 57
 comparative(s), 121–2; absence from
 Anatolian, 155
 comparative method, 3, 12; applied to cultural
 reconstruction, 14, 16
 comparative reconstruction, 3–4
 compensatory lengthening (induced by
 laryngeals), 58, 60–1
 complementizer position, 145
 composition and compounds, 77, 122; in
 Armenian, 346–7; in Classical Sanskrit,
 194; forms of numerals in, 123, 131–2;
 nominal, 122–3; verbal, 98
 conjugations (Italic), history of, 249
 conjunct verbs, 282, 286–8
 conjunctions, 134
 consonant clusters, 58–9, 63–4
 Constantine (Cyril), 373–4, 378
 constituents, syntactic, 137, 151
 Continental Celtic, 274; phonology and
 morphology, 278
 Čop, Bojan, 168–9; Čop's Law, 168
 Cornish, 296–7
 correspondence set, 4
 Cowgill, Warren, 133
 crane, 28
 creation myths, 27
 Cretan, 235
 Crimean Gothic, 312, 313
 Croatian, 377–8
 crow, 28
 culture, (P)IE *see under specific topics (e.g.*
myths)
 cuneiform, Mesopotamian, 159–61;
 limitations of, 162
 cuneiform, Old Persian, 212
 Cuneiform Luvian *see* Luvian
 Cuny, Albert, 76
 Curonian, 379
 curses, 25
 Cypriot, 223, 224
 Cyril *see* Constantine
 Cyrillic alphabet, 373
 Czech, 377
 Dacian, 404
 Daco-Mysian, 390, 406
 dactylic hexameter, 237
 Dafydd ap Gwilym, 293
damos, 379
 Danelaw, 320, 328
 Danish, 332
 Dardic (Indic group), 199
 Dardic (Nuristani), 180
 Darius I (the Great), 28, 210, 212, 213, 337
 dative, 105, 106–7, 114, 116
 daughter language, 4
 Dawn (goddess), 23
 dead languages, pronunciation of, 6
 debitive, 384
 declensions, Latin, 255
 deep structure, 137–8, 141, 150
 definite adjectives, 148, 305, 364, 367–8
 deities: Greek and Roman, 22; Proto-Indo-
 European, 22–4
 Delbrück, Berthold, 138
 demonstrative pronouns, 129–30
 Demotic Greek, 236
 denominatives, 90; in Italic, 249
 dental preterite, 306, 308
 dental-plus-dental clusters, 63, 181–2, 188,
 204, 275
 deponent verbs, 82, 251
 derivation, 76
 derivational suffixes: athematic, 110–13;
 thematic, 117–18
 derived verbs, 81
 desideratives, 91; in Celtic, 286
 determiner, 161
 deuterotonic, 286–8
 Devanāgarī, 150, 197, 355
devī-feminines, 119
 diathesis *see* voice
 digamma, 225, 237
 diphthongs, 61
 directive, 105, 156, 165
 discontinuous phrases, 139
 distraction, 139, 140, 239, 263
 distributive plurals, 156
 divine twins, 23–4
 Doric, 225
 dragons, Germanic, 26–7
 dragon-slaying myth, 26–7, 210
 Dravidian, 184, 188, 197

- dual, 105, 115; absence from Anatolian, 155
 ductus, 159
 Dukagjini, Leke, 392
 Dumézil, Georges, 17–18, 28
 Dutch, 327–8
 dwellings, PIE, 37

 Earth (goddess), 23
 East Baltic, 379
 East Germanic, 300, 311
 East Greek, 224, 229
 East Slavic, 375–6
 Eastern Armenian, 339
 Eddas, 329, 332
 Edgerton, Franklin, 65–6
eenie meenie miney mo, 292
 Eichner, Heiner, 109
 Elamite, 212
 enclitic(s) *see* clitic(s)
 endocentric compounds, 122
 Engadine, 260
 English *see* Middle English; Modern English;
 Old English
 Epic Greek, 223
 epics, Homeric *see* Homeric epics
 epics, Indian, 186
 ergative, 156, 157–8, 165
 Estonian, 379
 Etruscan, 246
 Etruscans, 245–6
 etymological method, 400–1
 Europe, Indo-Europeanization of, 43–4
 exchange, economic, 20
 existential verbs, syntax of, 144
 exocentric compounds, 122
ezāfe construction, 148, 213

 factitives, 90; in Hittite, 164; in Italic, 249
 Faliscan, 262
 fame, 29, 35, 239
 family, 18–19
 Family Books, 185
 Faroese, 331
 Farsi *see* Modern Persian
 Father Sky, 22–3, 24
 feminine, 118–20; absence from Anatolian,
 156; o-stems, 113
 Festus, Sextus Pompeius, 253
 finite verbs, prosody of in Sanskrit, 193
 Finnic peoples, 378
 fire: theft of (myth), 27; worship of, 24
 first-person pronoun, 127
 Flemish, 328
 folklore, 28
 formulae: legal, 26; poetic, 30, 223
 fosterage, 18–19
 foundation myths, 27
 Franconian, 327
 Frankish, 323
Fremdvokal, 353, 355
 French, 258–9, 323, 328
 fronting, 142, 144, 173
 full-grade, “wrong”, 117
 “functions” of PIE society, 17–18
 futhark, 310
 future, 91

 Gaelic, 283
Gaeltacht, 283
 Gaius (jurist), 21
 Galatians, 274, 277
 Gallego-Portuguese, 259
 Gamkrelidze, Tamaz V., 43, 54
 Gāndhārī, 196, 197
 Gathas, 205, 208
 Gaulish, 277–9
 Gauls, 277; society, 18
 Geg, 390, 391
 gender, 102–3; of o-stems, 113
 generative linguistics, 13, 137, 151
 genetic relationship, 3
 genitive, 104–5, 106, 114, 116; in Balto-
 Slavic, 234; in Greek, 234
 genitive-ablative, 104–5
 genre(s), textual, 13, 138
 Gepidic, 312
 Geraldus Cambrensis, 25
 German *see* Low German; Modern High
 German; Old High German
 Germanic, 300–9; contact with Balto-Slavic
 peoples, 300; heroic verse *see* poetry,
 Germanic; nouns and adjectives, 304–6;
 numerals, 364; phonology, 301–4;
 subgrouping of, 301; syntax, 309; verbs,
 306–9
 gift and counter-gift, 20
 Gimbutas, Marija, 41
 Glagolitic, 373
 glosses, Irish, 283
 glottalic theory, 54
 goats in PIE myth, 28
 ‘god’, words for, 22
 gods and goddesses *see* deities
 Gothic, 311–14; phonology and morphology,
 313; script, 312; syntax, 313
 Gothic Bible, 138
 Goths, 311–12
 gram, 37
 grammaticalization, 6
 grandson/nephew of the waters, 24
 Grassmann, Hermann, 188; Grassmann’s
 Law, 188, 227

- Great English Vowel Shift, 321
 Greek: alphabet, 225, 238, 245; Demotic, 236; dialects, 223–5 (*see also under individual names*); inscriptions, earliest, 245; labiovelars in, 227; noun, 234–5; phonology, 6, 226–31; syntax, 235; tribal divisions, 18; verb, 231–4
 “Greek Dark Ages”, 223
 Grimm, (Johann) Jacob, 9, 301; Grimm’s Law, 301–2
 Grimm, Wilhelm, 301
 griots, 29
 Grotefend, Georg Friedrich, 212
 Gruppenflexion, 357
 guest-friendship, 20
 Gujarati, 199
- Hale, Mark, 146
 Hallstatt culture, 274, 352
 hardening *see* Verschärfung
 Hattic people, 158
 Hattusas, 154
 hearth, 24, 37
 heaven, vault of, 23
 “heaviness” (syntactic), 151
 Hebrew, 324–5
 “*h₂e*-conjugation”, 163–4
 hellhound, 25
 Hengist and Horsa, 24
 Herakles, 27
 Herodotus, 246, 337, 379, 401, 404
 Hesychius, 236, 402, 404–5
 heteroclitic stems, 111
 “hic et nunc” particle, 85
h₁r-conjugation, 155, 157, 163
 Hieroglyphic Luvian *see* Luvian
 High German *see* Middle High German; Modern High German; Old High German
 High German Consonant Shift, 323–4, 326
 Hindi-Urdu, 199
 Hinduism, 205
 Hindustani *see* Hindi-Urdu
 Hispano-Celtic *see* Celtiberian
 Hittite, 158–67; discovery and decipherment, 154; importance for laryngeal theory, 76; law, 21; Luvianisms in, 168; nouns, 164–5; phonology, 6, 161–2; relative clauses, 165–6; stages of, 159; stress mobility, 162, 164; syntax, 165–6; verbs, 163–4; writing system, 159–61, 162
 Hittites, history of, 158–9
 Hoffmann, Karl, 111; “Hoffmann” suffix, 111
 holokinetic *see* amphikinetic
 homeland, PIE, 35–6, 39–43
 Homer, 7, 223, 404
- Homeric epics, 223, 237–8; syntax of proverbs in, 141
 Hopper, Paul, 54
 horse(s), 24, 37; sacrifice of, 24–5
 horseback riding, 41–3
 hospitality, 20
 Hrozný, Bedřich, 154
 Hübschmann, Heinrich, 338
 ‘human being’, words for, 22
 Hurrian, 337
 Hurrians, 184, 337
 hyperbaton, 139
 hysterokinetic, 107–8, 109
- ī (genitive sing.), 114, 247, 252, 277, 407
 Iberians, 278
 Icelandic, 331; Old *see* Old Norse
 ideograms, 160
 “ideology” (Dumézilean), 18
 IE *see* Indo-European family
 **ie/o*-presents, 89–90; in Italic, 249
 Iguvine Tables, 265
Iliad, 20, 29–30, 223
 Illyrian, 390, 400, 405–6, 407
 imperative, 95
 imperfect, 91
 “imperishable fame”, 29, 239
 “*i*-mutation”, 169
 inanimate *see* neuter
 India, settlement of, 183
 Indic, 183–4; *see also* Indo-Aryan languages; Middle Indic; Mitanni Indic; Old Indic; Sanskrit
 individualizing suffix, 305
 “Indo-Aryan”, 187
 Indo-Aryan languages, modern, 197, 199, 217; linguistic developments of, 198
 Indo-Aryans, migrations of, 183–4, 188
 Indo-European family: branches of, 8–9; establishment of, 8; relationship to other language families, 11–12; structure of, 8–9, 12; subgrouping of, 8–9, 12, 53–4, 91–2, 154–5, 181, 225–6, 247, 275, 338, 352, 403; *see also* Proto-Indo-European
 “Indo-Hittite”, 9–11, 155
 Indo-Iranian, 88, 180–3; clitics, 183; contacts with Slavs, 364; law, 19; morphology, 183; phonology, 181–3; relationship to other branches, 9, 181, 225–6, 247, 275, 338, 352, 403; syntax, 183; *see also* Indic; Iranian
 Indo-Iranians, 42–3
 Indra, 26, 28, 184
 Indus Valley Civilization, 184
 infinitives, 97
 infix, 75–6

- inflection, 76
 influence(s) (of one language on another) *see*
 borrowing(s)
 injunctive, 91, 133; in Sanskrit, 190–1
 instrumental, 105, 106–7, 114, 116
 Insular Celtic, 274; morphology, 281–2;
 phonology, 280–1
 interjections, 135
 internal derivation, 77, 110, 112, 120
 International Phonetic Alphabet, 48
 interpuncts, 150, 262
 interrogative/indefinite pronoun, 130
 Ionic, 223
 Iphitos, 27
 Iranian: geographical extent, 203; influence on
 Armenian, 337–8; influence on Slavic, 368;
 phonology, 203–4; *see also* Avestan;
 Middle Iranian; Modern Iranian
 Irish *see* Middle Irish; Modern Irish;
 Old Irish
 Irish law, 17, 20
i-stems, 103, 112–13
 Italian, 259
 Italic: linguistic unity, 246; nouns, 251–2;
 phonology, 248; verbs, 249–51
 “Italo-Celtic”, 9, 247, 275, 352
 Italy, prehistoric settlement of, 245
 iteratives (Hittite), 164
i-umlaut, 315–16, 318, 326, 329–30
 Ivanov, Vyacheslav V., 43, 54

 Jainism, 196
 Jakobson, Roman, 54
 Japanese, 160, 197
 Jasanoff, Jay, 164
 Jones, Sir William, 8, 13, 185
 Jupiter, 28
 justice, public enforcement of, 22
 Jutes, 316

 Kafiri *see* Nuristani
 Kālidāsa, 186
 Kanesh, 158
 Kannaḍa, 197
 Kashmiri, 199
 Kashubian, 377
 Katharevusa, 236
 Katz, Joshua T., 359, 361
kavya, 186
 kennings, 329
 Kentish, 317
 Kharoṣṭhī alphabet, 197
 Khotanese, 197, 216, 351
 Khvalynsk culture, 43
 king, 19
 kingship ritual, 24

 kinship, 18; kinship terms, 18, 112
 Knudtzon, Jorgen, 154
 Koine, 235–6
 Kurdish, 217
 kurgan culture, 41
 kurgans, 41, 42
 Kuryłowicz, Jerzy, 76, 289; Kuryłowicz’s
 Fourth Law of Analogy, 289
 Kushans, 216

l/n-stems, 111
 Laconian, 236
 Ladin, 260
 Langobardic, 323
langue d’oc, 258
langue d’oïl, 258
 Laroche, Emmanuel, 159–61
 laryngeal hiatus, 189, 204, 208
 laryngeals, 56–8, 75–6; in Albanian, 394; in
 Armenian, 42, 229; in Balto-Slavic, 366–7;
 in Celtic, 276; coloring, 57; consonantal
 reflexes of, 56, 155, 156, 173, 342;
 evidence for, 71, 75–6; in Germanic,
 303–4; in Greek, 228–9; in Indo-Iranian,
 182, 204; loss of, 57; phonetic value of,
 58; in Phrygian, 229; roots containing,
 71–2; in Tocharian, 354–5; vocalized, 57
 Late Cornish, 296–7
 Latin, 252–8, 260–2; earliest remains, 252;
 nouns, 255–6; phonology, 6, 253–4; verbs,
 82, 83, 256–7; word order in, 145
 Latino-Faliscan, 246; phonology, 252
 Latvian, 379, 383–5
 law: Albanian, 392; Hittite, 21; Indo-Iranian,
 19; Irish, 17, 20; PIE, 21, 186; Roman, 19,
 21
 Law of Open Syllables, 368–9
 Law of the Palatals, 181, 183
 lawsuits, 22
 Lemnian, 246
 lenition: Celtic, 281; Luvian, 169; Old Irish,
 284
 Lepontic, 277
 Lesbian, 223
 Lethe, 25
 Letoon Trilingual, 172–3, 174
 Lie, religious, 28
 Lindeman, Fredrik Otto, 66; Lindeman’s Law,
 65–6; Lindeman’s Law variants, 66
 Linear A, 222
 Linear B, 13, 222
 Lithuanian, 379, 381–3
 Llywarch Hen, 29
 loanwords *see* borrowing(s)
 locative, 103, 105, 106, 116, 381–3;
 endingless, 105

- long syllabic resonants, 189; in Celtic, 276; in Greek, 229
 long-vowel preterites, 93, 396
 Low German, 326
 Lug, 28
 Lusitanian, 278, 408–9
 Luther, Martin, 324
 Luvian, 167–71; connection with Troy, 168; Cuneiform, 167–8; Hieroglyphic, 167–8; hieroglyphic script, 167–8; influence on Hittite, 168; morphology, 169–70; phonology, 168–9
 Lycian, 172–4; alphabet, 173; phonology, 173; varieties of, 173; verb placement, 142, 173; word order, 144
 Lydia, 175
 Lydian, 175–6

 *-*m*- (case element), 107, 156, 364
 Macedonian (ancient Balkan lang.), 400, 404–5
 Macedonian (Slavic lang.), 378
 Macha, 24
 magic, 25; spells, 325
Mahābhārata, 186
 Malayalam, 197
 Mallory, J. P., 36
 Mani, 215
 Manichean Middle Persian, 216
 Manicheism, 215
Männerbund, 19, 328
 Mannus and Tuisto, 27
 Manu, 27
 manuscript transmission, 7
 Manx, 290–1
 Marathi, 199
 marriage, 18
 Marrucinian, 269
 Mars, 28
 Marsian, 269
 Mazdayasman religion *see* Zoroastrianism
 McCone, Kim, 19
 Medes, 337
 Median, 212
 mediopassive, 83; participle, 97–8
 Meillet, Antoine, 338
 Melchert, Craig, 60, 174
 Mercian, 316
 merism, 33
 mesostatic, 107
 Mesrob, 338
 Messapic, 407–8
 meter, poetic *see* scansion
 Methodius, 373–4, 378
 metrics, 31, 151
 *-*mi*-, added to *-*(o)h2*, 89–90

mi-conjugation, 157, 163
 middle: meaning of, 82–3; relationship to perfect, 93
 Middle Armenian, 339
 Middle Breton, 295
 Middle Cornish, 296, 297
 Middle Dutch, 327–8
 Middle English, 258, 320–1
 Middle High German, 324
 Middle Indic, 189, 196
 Middle Iranian: developments, 214–15; scripts, 215
 Middle Irish, 283
 middle-passive *see* mediopassive
 Middle Persian, 215–16
 middle verbs, 82
 Middle Welsh, 293
 Milyan, 173
 Mimir, 25
 misspellings, 7
 Mitanni documents, 28
 Mitanni empire, 184
 Mitanni Indic, 183, 184, 190
 Mitra, 28, 184
 Mjöllnir, 23
 mobile accent system: evidence for and fate of, 164, 190, 217, 230–1, 248, 254, 301, 303, 366, 376, 377, 381, 383, 394; in PIE, 62
 Modern Armenian, 339–40
 Modern English, 321
 Modern High German, 324
 Modern Iranian, 217–18
 Modern Irish, 283
 Modern Persian, 213, 217
 Modern Slavic languages, 375–8
 Modern Welsh, 293
 Möller, Hermann, 57, 76
 Mongolian languages, 352
 mood, 83, 95–6
 morphological change, 69
 Mother Earth, 23
 movement, syntactic, 138, 144
 Movsēs Xorenac'i, 338–9, 347
 Mozarabic, 259
 MQ texts, 351
 multiplicative adverbs, 133
 mummies, Central Asian, 352
 mutations, 280–1
 Mycenaean, 222–3, 227, 232
 mythology, Germanic, 329
 myths, PIE 26–8

 *-*n*-, loss of in PIE, 64
 name-giving ritual, 35
 names: divine, 22–4; personal, 34–5
 naming practices, 34–5

- Narten, Johanna, 88
 "Narten" ablaut, 88, 92
 Narten presents, 88, 93, 96
 nasal presents, 75–6, 88; in Hittite, 164
 nasalization, 281, 284
 nasalized vowels (Slavic), 370
 Nasatyas, 23, 28, 184
 Nechtan, 24
 negation, 133, 331; syntax of, 149
néktar, 29
 Neogrammarian hypothesis, 5
 "neo-lengthened grade", 367–8
 Neolithic, 11, 38
 Nepali, 199
 Neprune, 24
 neuter, 103
 neuter plural, syntax of, 118, 143, 235
 New Phrygian, 402
 ninth class of Indo-Iranian verbs, 88, 285
 *-*nó-* in Wakhi, 218
 nominal sentences, 143
 nominative du., 104, 105, 106, 114, 115
 non-ablauting roots, 74
 non-finite verbal formations, 96–8
 non-material culture, 16
 Norman French, 258, 320, 338
 Normans, 328
 Norse *see* Old Norse
 North Germanic, 300
 "North Oscan", 269
 Northumbrian, 316
 Norwegian, 331–2
 Nostratic, 12
 Notker Labeo, 150, 151
 noun phrases, 139
*n-*stems, 111; in Germanic, 304–5
*nt-*participle, 97; in Anatolian, 97, 155, 164
 number, 102
 numerals, 130–2; cardinal, 131; feminine of, 120, 132; ordinal, 122, 132; in Romance, 1–3; in Slavic, 373
 Nuristani, 180

 *-*ō-* (possessive suff.), 116–17
 oak tree, relation to thunder god, 23
 Occitan, 258
 October Equus, 24–5
 Odin, 28
Odyssey, 20, 223
 Ogami, 282
 o-grade presents, 91
 Old Albanian, 392
 Old Avestan, 204–5, 207, 209; *see also* Avestan
 Old Breton, 294–5
 Old Bulgarian, 374
 Old Church Slavonic, 373–5; scripts, 373
 Old Cornish, 296
 Old English, 315, 316–20; morphology, 319–20; phonology, 317–18
 "Old European" culture, 41
 Old French, 258
 Old Frisian, 328
 Old Geg, 391
 Old High German, 315, 322–4; dialects, 323; phonology, 323–4
 Old Icelandic, 329
 Old Indic, 184–5
 Old Irish, 282; law, 17; noun, 288–9; phonology, 284–5; verb, 285–8
 Old Lithuanian, 381
 Old Low Franconian, 327
 Old Norse, 328; influence on English, 328–9; morphology, 330–1; phonology, 329–30
 Old Norwegian, 329
 Old Persian: morphology, 213; phonology, 213; script, 212; where and when spoken, 210–12
 Old Phrygian, 402
 Old Prussian, 379, 385–7
 Old Russian, 375–6
 Old Saxon, 315, 323, 325–6
 Old Swedish, 329
 Old Tosk, 391
 Old Welsh, 291, 293–4
 Omaha system, 18
 online resources, 14, 177
 optative, 96; absence from Anatolian, 155
 oral-formulaic poetry, 29–30, 237
 Oriya, 199
 Orm, 150, 320–1
 Oscan, 246, 263, 268–9; phonology and scripts, 268
 Osco-Umbrian *see* Sabellic
 Ossetic, 217
*o-*stems, 113
 Osthoff, Hermann, 64; Osthoff's Law, 64
 Ostrogoths, 311–12
 'overcome death', 27, 29

 Paelignian, 269
 Pahlavi, 216; *see also* "Book Pahlavi"
 Palaic, phonology, 171
 palatalization: in Irish, 284; in Latvian, 384; in Old English, 217; in Slavic, 369; in Tocharian, 354; in Umbrian, 265
 paleography, 159
 Pāli, 196–7
 Pamphylian, 224
 Papini, 186
 Parsis, 204
 Parthian, 215, 337

- participles, 97–8
 particle(s) *see* clitic(s)
 Pashto, 208, 217
 passive, 83, 89–90
 patron (of poet), 29, 209
 patronymics, 121, 347
 “P-Celtic”, 280, 291
 perfect, 93–5; absence from Anatolian, 155;
 meaning of, 94–5, 192; participle, 98;
 relationship to middle, 93
 Perunas, 23, 28
 Persian *see* Old Persian; Middle Persian;
 Modern Persian
 Persian Empire, 210, 215, 337
 personal endings, 83–4
 personal pronouns, 126–9; archaic nature of,
 126; enclitic, 128–9, 147; infixed, 282
 Perun, 23
 philology, 7–8, 13–14, 237–8
 Phoenicians, 225
 phonetic complement, 160
 phonology–syntax interface, 149, 151
 Phrygian, 401–4
 Phrygians, 337, 401–2
 Pictish, 291
 PIE *see* Proto-Indo-European
 Pindar, 225
 Pisidian, 176
 Plautus, 6, 33, 145, 252
 pluperfect, 81
 plural (nouns), 105–7, 115–16
 plurative, 102, 357
 poet, role of, 29; and patron, 29, 209
 poetic grammar, 138
 poetic “licenses”, 138, 237
 poetics, 29–34
 poetry, Germanic, 33, 310, 329; *see also*
 poetics
 Polabian, 377
 Polish, 376–7
 Pomeranian, 377
 Pontic-Caspian steppes, 41, 42–3
 Portuguese, 259
 positive used as superlative, 122
 possession, formations indicating, 111, 112,
 113, 116–17, 121, 122–3
 postpositional phrases, 139
 postpositions, 133–4
 postpositive clitics, 134; types of, 146–7
 Prakrits, 196
 prefixation, 77
 prefixes, 123
 pre-Greek substrate languages, 222
 prepositional phrases, 139
 prepositions, 133–4; conjugated, 282
 “Pre-Samnite”, 269
 present stem, 87
 preterite-presents, 308
 preverbs, 98–9, 139–40, 304
 primary cases (Tocharian), 356–8
 primary endings, 84, 95; in Italic, 251, 264
 primary *-jelo-presents, 89–90
 primary verbs, 81
 primeval being, 27
 Primitive Irish, 282–3
 Primitive Welsh, 293
 privative prefix, 133, 149
 productivity (morphological), 5, 69
 Prometheus, 27
 pronominal adjectives, 130
 pronominal declension, 129; influence on
 nouns, 115, 234
 pronouns *see under individual types*
 property and wealth, 19
 prosodic domains, 149
 prosody (of phrases and sentence), 149–51;
 reconstruction of, 151–2
 proterokinetic, 107–8, 108–9, 208, 375
 proterotonic, 286–8
 “prothetic” vowels, 71, 229
 proto-culture, 16
 Proto-Indo-European, 9, 14; dating of, 36,
 38–9, 43; homeland, 35–6, 39–40, 41,
 42–4; society, structure of, 17, 19; *see also*
 under specific topics (e.g. ablative, kinship)
 Proto-Indo-Europeans, 14; expansion of, 37
 proto-language, 4
 Proto-Romance, 4
 Provençal, 259
 proverbs in Homer, syntax of, 141
 Prussian *see* Old Prussian
 Psammetichus, 401
 Ptolemy, 291, 379
 Punjabi, 199
 Puruṣa, 27
 Pushan, 28

 q (Mycenaean), 223, 227
 “Q-Celtic”, 280
 quantitative verse, 31
 Quirinus, 28

 *-r (middle marker), 86, 251, 359
 Raetic, 246
 Rāmāyaṇa, 186
 Rask, Rasmus Krstian, 301
 raven, 28
 reciprocity, 20, 29, 209
 reconstruction, linguistic, 3–4, 12–13, 17, 48,
 69–70, 72, 126–7; *see also* syntactic
 reconstruction
 reduplicated aorist, 92–3

- reduplicated nouns, 112, 117
- reduplicated presents, 88, 90
- reduplication, 77
- reflexive adjective, syntax of, 143
- reflexive pronoun, 130
- regularity of sound correspondences, 4
- relational adjectives, 169, 176
- relative clauses, syntax of, 147–8; in Hittite, 165–6
- relative pronoun, 114, 130; in Carian, 148; in Celtic, 148, 277; in Hittite, 148, 165–6; in Persian, 148, 213; reanalysis of, 148
- religion, Baltic, 23
- Renfrew, Colin, 43
- restitution, law of, 21
- Rhaeto-Romance, 260
- rhotacism, 248, 253, 265–6, 310, 315
- Rig Veda, 7, 18, 30–2, 42, 83, 138, 150, 184, 185, 189, 195, 196, 205, 208, 209
- r*-initial words, 156
- ritual, 24–5
- Rix, Helmut, 109
- r/n*-stems, 111; in Anatolian, 155, 165; in Armenian, 344; in Germanic, 305; in Italic, 251
- Roman alphabet, 225
- Roman law, 19, 21
- Romance languages, 196, 257, 258–60
- Romani, 199
- Romanian, 260
- Romansh, 260
- Romulus and Remus, 27
- root, definition of, 69
- root aorist, 92
- root athematic presents, 87
- root athematic verbs, ablaut in Hittite, 163
- root extensions, 72
- root nouns, 108, 109
- root structure, 70; constraints on, 72
- roots with *a*, 72; ablaut of, 75
- r*-stems, 111
- Rugian, 312
- “ruki”, 182, 189, 364, 365–6, 369, 379
- runes, 309–10
- Runic, 301, 310, 328, 329
- Rus’, 328, 375, 376
- Russian, 375–6
- Russian Church Slavonic, 375
- *s*, voicing of, 55, 63
- Sabellic, 246, 264
- Sabine, 269
- sagas, 329
- Saka, 216
- salmon, 40
- Samnites, 246
- sandhi, 150, 190; in Avestan, 208; in Celtic, 281; in Old Persian, 212; in Sanskrit, 190
- Sanskrit, 75–6, 83, 184–96; Classical, 186; compounding, 194; definition of, 184–5; nouns, 193–4; phonology, 6, 187–90; pronouns, 194; retroflexion, 188; syntax, 194; verbs, 191–3
- s*-aorist, 92
- Sardinian, 248, 259
- Sarmatian, 217
- Sassanian dynasty, 205, 215
- satem, 53–4; in Balto-Slavic, 365–7
- Saturnian verse, 33
- Saussure, Ferdinand de, 56–7, 75–6
- Saxons, 316
- Scandinavian, 328, 331–2
- scansion, 7, 31–3, 106, 151, 228
- Schindler, Jochem, 65–6, 109, 110
- Schleicher, August, 48
- “schwa primum”, 61–2; “schwa secundum”, 61–2
- scientific method, 17
- Scottish Gaelic, 290
- scriptio plena, 161
- scripts, 167–8, 190, 206, 277; Indian, 197; Southeast Asian, 198
- Scythian, 203
- *-s*-desiderative/future, in Italic, 250
- second compound member, ablaut of, 123
- second-person pronoun, 128
- secondary cases (Tocharian), 356–8
- secondary endings, 84; in Italic, 251, 264
- secondary verbs, 81
- semi-deponent verbs, 83
- Semitic, 57, 225, 324–5; alphabet, 197
- sentence-connecting particles, 135, 286
- sentence-connective clitics, 146–7
- sentence connectives (Anatolian), 157
- sentential clitics, 147
- separable prefixes (German and Dutch), 140
- Serbian, 377–8
- Serbo-Croatian, 377–8
- Sicel, 269
- Sidetic, 176
- Sievers, Eduard, 65–6; Sievers’ Law, 65–6
- sigmatic aorist, 92
- silver, 39
- similarities between languages, causes of, 1
- simple thematic presents, 89–90; absence from Anatolian, 155
- Sindhi, 199
- singular (nouns), 104–5, 113–14
- singulative, 295
- Sinhalese, 197, 199
- skaldic poetry, 329
- *skelo*-presents, 90; in Hittite, 164

- sky-god *see* Father Sky
 Slavic, 368–73; and Germanic, 300, 368; and
 Iranian, 203, 368; nouns, 371–2;
 phonology, 368–72; verbs, 371
 Slavs, prehistory, 368
 Slovak, 377
 Slovenian, 377
 Slovincian, 377
 *-sm- (pronominal stem formant), 129, 385
 s-mobile roots, 70–1
 Sogdian, 216–17
 Soldier's Oath (Hittite), 25
 soma, 26–7
 Sorbian, 377
 soul, journey of, 25
 South Picene, 267
 Spanish, 259
 “special clitics”, 147
 spelling, 321
 spells, 25, 28
 split ergativity, 197, 217
 “spurious diphthongs”, 226, 230
 Sredny Stog culture, 43
 *ss, simplification of in PIE, 63
 s-stems, 112
 Stang, Christian, 63–4; Stang's Law, 63–4
 stative (meaning of perfect), 95, 192
 statives, 90–1, 123; in Italic, 249
 stød, 332
 stress in PIE, 62; *see also* mobile accent
 system
 “strong” ablaut grade in verbs, 87
 “strong” adjectives in Germanic, 305
 strong cases, 103
 “strong” verbs: Germanic, 306–8; Celtic, 285
 strophic style, 32–3
 Sturluson, Snorri, 329
 Sturtevant, Edgar, 9, 155, 162; Sturtevant's
 Law, 162
 Styx, 25
 subgrouping, 8–9, 12, 53–4, 91–2, 154–5,
 181, 225–6, 247, 275, 338, 352, 403
 subject pronouns, overt, 143
 subject–verb agreement, 118, 143
 subjunctive, 95; absence from Anatolian, 155;
 meaning of, 96
 subordinate clauses, syntax of, 147–8
 subordinating conjunctions, 134
 Sumerian, 160
 Sumerograms, 160
 Sun (god), 23
 superlative, 122, 132; expressed by positive,
 122
 supine, 97
 suppletion, 81, 126
 surface structure, 138, 141
 Swedish, 329, 332
 syllabaries, 13
 syllabic resonants, long *see* long syllabic
 resonants
 syllabification, 64–5, 354–5
 syllabograms, 160
 sympathetic magic, 166
 syntactic reconstruction, methodology of,
 140–1; problems with, 137–8
 syntax: hierarchical organization of, 137;
 interaction with phonology, 149–51
 Szemerényi, Oswald, 64; Szemerényi's Law,
 64, 104, 108
 -z in 3 sg. middle, 86–7, 192
 taboo deformation and replacement, 28
 Tacitus, 27, 300, 316, 379
 Tajiki, 217
 Tamil, 197
 Tarhunnaš 26–7
 Tartessian, 278
 teleutostatic, 107
 Telugu, 197
 tense, 81
tertium comparationis, 8
 *teutā, 19
 Teutonic *see* Germanic
 texts *see* genre(s), textual
 textual transmission, 7
 thematic aorist, 92
 thematic inflection, 77–8; nouns, 113–16;
 presents, 89–90
 thematic vowel, 77; in forming ordinals, 132
 Thor, 23, 28
 “thorn” clusters, 59–60, 71, 354, 402; in
 Indo-Iranian, 182; in Pāli, 197
 Thracian, 390, 400, 404
 Thracians, 401
 “Thraco-Illyrian”, 406
 “Thraco-Phrygian”, 404
 thunder and lightning, god of, 23
 Tibetan, 197
 tmesis, 140
 *-tō-verbal adjective, 98; absence from
 Anatolian, 155
 Tocharian, 197; nouns and pronouns, 356–8;
 phonology, 353–6; script, 353; verbs,
 358–9
 Tocharian A, 351
 Tocharian B, 351
 Tocharians, origins of, 351
tomós/tómos nouns, 116
 tones in Scandinavian, 332
 “tool” nouns, 118
 topicalization, 144
 “ToRT” groups, 370–1

- Tosk, 390, 391
t-preterite, 281–2
 translation literature, 138, 309
 trees and plants, 40
 trifunctionalism, Dumézilean, 17–18, 28
 “triple reflex”, 57, 229
 Troy, 167, 168
 Truth, religious, 28, 31
 truth, ruler’s, 289
 Tsakonian, 236
tudāti-presents, 89–90
 Tumshuqese, 216, 351
 Turkic languages, 339, 352, 357
 Twin (primeval being), 27
 twins, divine, 23–4
 typology, linguistic, 54

ʔ-*u*- (adj. suff.), 123
 Ukrainian, 376
 Ulfilas *see* Wulfila
 Umbrian, 263, 265–7
 Umbrians, 246
 umlaut, 301, 313, 315–16, 318–19,
 329–30; *see also* *i*-umlaut; *u*-umlaut;
 w-umlaut
 unaccusative verbs, 82
 universals, linguistic, 1
 Uralic languages, 352, 357
 Urartians, 337
 Urartu, 337
 Urdu, 199
u-stems, 103, 112–13
u-umlaut, 329–30

 Vahagn, 24
 Vandalic, 312
 Varuna, 28, 184
vāyav indraś ca construction, 140
 Vedic literature, 185–6
 Vedic Sanskrit *see* Sanskrit
 vehicles, wheeled, 36, 38
 Veneti, 406
 Venetic, 246, 406–7
 Ventris, Michael, 222
 Venus, 22
 verb-final word order, 142
 verbs, deponent, 82, 251; derived, 81; middle,
 82; primary, 81; prosodic status of, 99,
 151, 231; semi-deponent, 83; tense-stems,
 81; unaccusative, 82
 Verner, Karl, 303; Verner’s Law, 301, 303,
 307; Verner’s variants, 307, 313, 319–20
 Verschärfung, 312, 313, 329–30
 verse, quantitative, 31
 Vesta, 24
 Vestinian, 269

 Vikings, 291, 328
 Villanovan culture, 246
 Visigoths, 312
 vocalized laryngeals, 71; in Anatolian, 173
 vocative, 104, 105, 106, 114, 115; prosody of
 in Sanskrit, 193; syntax of, 140
 voice, 82
 voiceless aspirates, 188
 voicing assimilation, 63, 182
 Volsian, 269
 vowel affection, 281, 285
 vrddhi-derivatives, 116–17
vṛkī-feminines, 119
Vṛtra, 26
 Vulgar Latin, 196, 253, 257–8

 Wackernagel, Jacob, 138, 146; Wackernagel’s
 Law, 138, 146–7, 183, 235, 313
 Wakhi, 216, 218
 warrior class and culture, 17, 29, 35, 328
 Watkins, Calvert, 26–7, 32
 “weak” ablaut grade in verbs, 87
 “weak” adjectives in Germanic, 305
 weak cases, 103
 “weak” declension of German nouns, 305
 “weak” verbs: Germanic, 306, 308; Celtic,
 285
 Welsh, 291, 292–4; phonology, 291–2
 West Baltic, 379
 West Germanic, 300–1, 315; phonology,
 315–16
 West Greek, 224–5
 West Saxon, 316
 West Slavic, 376–7
 Western Armenian, 339–40
 wheel, 36, 38
 WH-movement, 145
 Wilusa, 168
 wine, 38
 Winter, Werner, 365; Winter’s Law, 365
 witness (legal), 22
 Woden, 22
 wolf, 19, 28
 word-final voicing neutralization, 64, 114
 word-level clitics, 146
 word order, 141–2
 word structure, 76–7
 writing systems, 7, 12–13, 150, 159–61, 173,
 309–10
 Wulfila, 6, 309, 312
w-umlaut, 329–30

 Xanthos Stele, 172

 Yaghnobi, 217, 218
 Yama, 27

Yamna culture, 41–3

yers, 369–70

Yiddish, 324–5

Ymir, 27

Younger Avestan, 204–5, 210, 212; *see also*
Avestan

Zarathuštra, 205

“Zend”, 204

zero-grade, formation of, 74

zero-suffix, 77, 108

Zoroaster *see* Zarathuštra

Zoroastrianism, 24, 28, 204–5, 338